HETEROCYCLIC SMALL MOLECULE PEPTIDOMIMETICS

A Dissertation

by

JING LIU

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2008

Major Subject: Chemistry

HETEROCYCLIC SMALL MOLECULE PEPTIDOMIMETICS

A Dissertation

by

JING LIU

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Kevin Burgess
Committee Members,	Brian T. Connell
	Rajesh Miranda
	Eric E. Simanek
Head of Department,	David Russell

December 2008

Major Subject: Chemistry

ABSTRACT

Heterocyclic Small Molecule Peptidomimetics.

(December 2008)

Jing Liu, B.S., Nankai University

Chair of Advisory Committee: Dr. Kevin Burgess

Polymer-supported synthesis of a close analog (*i.e.* **A)** of an early lead, a 14-membered ring peptidomimetic **D3**, was described. The monovalent molecule was attached to different length linkers, and they were then paired sequentially on a triazine scaffold via our previously published methodology to give a small library of bivalent compounds **1** representing all combinations of linkers of the different lengths in a fast and efficient combinatorial manner. Cellular assays identified **1-ss** as a TrkA receptor antagonist towards NGF and it was shown to bind TrkA with ~200 nM affinity and retains high selectivity towards TrkA in binding assays.

A set of monovalent diketopiperazine (DKP) mimics **4-7** was synthesized efficiently from corresponding dipeptides via intramolecular S_N2 cyclization reactions in solution. These DKP compounds contain two amino acid side-chain functionalities to mimic the sequences that occur at "hot-spots" in loop regions. The monovalent mimics were assembled into a library of biotin-labeled bivalent molecules **9** via the combinatorial strategy described above with some modification. In primary screening, compound **9gg** showed preferential binding to TrkC receptors in FACScan assay and blocked the trophic activity of NT-3 in TrkC cells at 10 uM in cell survival assay.

The preparation of monovalent 1,3,4-oxadizole-based mimics **12** was achieved from corresponding amino acid building blocks on gram scale in a highly efficient solution phase parallel synthesis manner in good yields. These heterocyclic compounds feature various natural amino acid side-chain functionalities including those occuring most frequently at hot-spots such as those of Tyr, Lys, Glu and Ser. Attempts to assemble them into bivalent molecules were done by coupling the monovalent mimics to

the triazine scaffold sequentially in solution and simply manipulating the solvent systems. For some reasons, some reactions did not proceed cleanly. Studies have been carried out and the problems were partially solved. The biological activities of these oxadiazoles are under investigation. So far, six compounds have shown activities in four different bioassays.

Two different peptidomimetic types that resemble protein A and protein G binding regions were generated and tested as binding factors in affinity columns for purification of IgG. They are cyclic hexapeptides **19**, which were prepared via Fmoc-SPPS and solution phase intramoleculer macrocyclization, and heterocycle-based small molecules **22** and **23** featuring a variety of aromatic functionalities generated via solution phase parallel synthesis. Four compounds showed some affinity towards a Fab fragment of IgG in SAR screening, and they were attached to a dendrimer core on a solid support to give four multivalent mimetics **25**.

DEDICATION

To my country and my parents

ACKNOWLEDGEMENTS

I would like to express my gratitude to my advisor, Dr. Kevin Burgess for his enthusiasm and guidance during my research in his group. I also would like to thank the members of my committee, Drs. Brian Connell, Eric Simanek and Rajesh Miranda; and Drs. Hagan Bayley, Gary Sulikowski, Victoria DeRose, and Andy LiWang, who used to be on my committee. I also want to thank Dr. Jiong Yang for his substitution for Dr. Simanek at my final defense.

I would like to acknowledge all my colleagues especially former group members Drs. Yu Li Angell, David Chen, Genliang Lu, Zebin Xia (for doing molecular modeling in chapter V), Sam Reyes and Mookda Pattarawarapan for their friendship and helpful discussions in doing my research, and to Dr. Andrey Malakhov for cooperation in D3 dimerization project, as well as to all the current members of the Burgess Lab for their companionship. I am also very grateful to Dr. Amber Schaefer for proofreading my entire dissertation.

Finally, I would like to pay special thanks to my family and my boyfriend, Dr. Qingsong Liu for their love, support and patience.

TABLE OF CONTENTS

Page

ABSTRACT iii

DEDICATION v

ACKNOWLEDGEMENTS vi

TABLE OF CONTENTS vii

LIST OF FIGURES x

LIST OF TABLES xiv

LIST OF SCHEMES xv

CHAPTER I INTRODUCTION 1

1.1 Heterocyclic Amide Bond Surrogates of Dipeptide Mimetics 1
1.2 Five-membered Heterocyclic Dipeptide Mimetics 4
1.2.1 Azole-based Dipeptide Mimetics 4
1.2.2 Other Five-membered Heterocycle-based Dipeptide Mimetics 10
1.3 Six-membered Heterocyclic Dipeptide Mimetics 14
1.4 Larger (>6) Heterocyclic Ring Systems as Dipeptide Mimetics 16
1.5 Other Heterocylic Dipeptide Mimetics 19
1.6 Monovalent Heterocyclic Dipeptide Mimetics Designed in Our Group 20
1.7 Conclusion 23

CHAPTER II DIMERIZATION OF A MONOVALENT SMALL MOLECULE LIGAND CHANGES A TrkA RECEPTOR AGONIST INTO AN ANTAGONIST 24

2.1 Introduction 24
2.2 Solid Phase Synthesis of the Monovalent Cyclic Peptidomimetic A 31
2.3 Synthesis of the Labeled Bivalent Peptidimimetics 1 with s/m/l Linkers 32
2.4 Binding Assays 36
2.4.1 Direct Binding Studies with Labeled Peptidomimetics 36
2.4.2 NGF Binding Competitive Studies 37

Page

2.5 Receptor Dimerization Studies 40
2.6 Cellular Assays 42
2.6.1 Cell Survival Assays 42
2.6.2 Antagonism of Biochemical Signals by 1-ss 46
2.7 Docking/Modeling 47
2.8 Discussion and Summary 50

CHAPTER III A COMBINATORIAL METHOD TO SOLUTION-PHASE SYNTHESIS OF BIOTIN-LABELED BIVALENT BETA-TURN MIMICS 53

3.1 Introduction 53
3.2 Preparation of Monovalent Diketopiperazine Peptidomimetics 56
3.3 Application of the Solution-phase Procedure to Make a Library of Biotin-labeled Bivalent Mimics 60
3.4 Biological Assays 64
3.5 Summary 68

CHAPTER IV SOLUTION PHASE SYNTHESIS OF OXADIAZOLE-BASED BETA-TURN MIMETICS AND ATTEMPTS TO ASSEMBLE THEM INTO BIVALENT MOLECULES 70

4.1 Introduction 70
4.2 Preparation of 1,3,4-Oxadiazole Based Beta-Turn Mimetics 74
4.3 Preparation of Bivalent Compounds 78
4.4 Biological Assays 83
4.4.1 Screen for Chemicals that Inhibit the RAM Network 83
4.4.2 Primary Cell-based High Throughput Screening Assay for Inhibitors of Wee1 Degradation 84
4.4.3 uHTS Identification of Diaphorase Inhibitors and Chemcical Oxidizers: Counter Screen for Diaphorase-based Primary Assays 84
4.4.4 uHTS of Mcl-1/Bid interaction inhibitors 85
4.5 Summary 86

CHAPTER V SYNTHESIS OF CYCLIC PEPTIDOMIMETICS OF PROTEIN A AND SMALL MOLECULE MIMETICS OF PROTEIN G FOR PURIFICATION OF IMMUNOGLOBLIN G VIA AFFINITY CHROMATOGRAPHY 87

5.1 Introduction 87
5.2 Design and Synthesis of Cyclic Peptidomimetics of Protein A 91
5.2.1 Rationale 91
5.2.2 Syntheses of Cyclic Peptidomimetics 19 of Protein A 95

Page

5.3 Design and Synthesis of Small Molecule Mimetics of Protein G.... 98
5.3.1 Rationale 98
5.3.2 Synthesis of Small Molecule Mimetics 22 and 23 of Protein G 101
5.3.3 SAR Screening of Small Molecular Protein G Mimetics 22 and 23 104
5.3.4 Synthesis of Multivalent Mimetics 25 from Active Monovalent Ligands 22 106
5.4 Summary 109

CHAPTER VI CONCLUSIONS 111

REFERENCES 114

APPENDIX A 126

APPENDIX B 131

APPENDIX C 202

APPENDIX D 275

VITA 305

LIST OF FIGURES

Page

Figure 1.1. Schematic view of dipeptide mimetics based on a heterocyclic amide bond surrogate featuring two amino acid side-chains.....2

Figure 1.2. Various sugar amino acids as a. dipeptide isosteres; and b. Gly-Gly replacement in Leu-enkephalin analogs.....3

Figure 1.3. Various conformationally constrained bicyclic dipeptide mimetics.....3

Figure 1.4. Three types of pseudopeptides as Phe-Gly replacements. 5

Figure 1.5. Xaa-1,5-triazole-Ala and Xaa-1,4-triazole-Ala modules as mimics of Xaa-*cis*-Pro segments.....6

Figure 1.6. Azole-based Leu-Trp and Trp-X surrogates in potential endothelin antagonists. 7

Figure 1.7. Dipeptide mimetics of calpain based on 2,5-diketopiperazines. 14

Figure 1.8. 1,4-Oxazepines as conformationally constrained *gauche* (-) Phe-X dipeptide mimetics.....17

Figure 1.9. Nine-membered-ring lactam-based dipeptide mimetic as β-turn analog.....19

Figure 1.10. Various heterocyclic dipeptide mimetics. a. dipeptide mimetics mimicking particular dipeptide sequences; b. dipeptide mimetics mimicking secondary structures; c. conformationlly restricted dipeptide mimetics. 19

Figure 1.11. a. The key distance of C^{β}-separations of the $i + 1$ to $i + 2$ residues of a type 1 β-turn and of the triazole-based turn mimics featured here; b. an overlay of the mimics onto a type 1 β-turn. 21

Figure 2.1. Some β-turn nomenclature and some cyclic peptidomimics with ring-fused C^{10} motif.....25

Figure 2.2. Previously identified lead compounds D3 and MPT18. 26

Page

Figure 2.3. a. Generic scheme of the monovalent mimic A was attached to a short s, a medium m, and a long linker l bearing a nucleophile, then paired on a triazine scaffold to which either a FITC or a TEG tag was incorporated; b. structure of the bivalent compounds 1-ss, 1-sm, 1-mm, 1-sl, 1-ml, and 1-ll. 28

Figure 2.4. Direct binding FACScan assay with FITC-peptidomimetics. Cells expressing the indicated receptor were bound with test ligand (20 μM) at 4°C. After washing data was acquired and analyzed by FACScan/ CellQuest with background subtracted. Mean channel fluorescence ± sem, n = 3-6 independent experiments..... 37

Figure 2.5. Binding competition studies using NIH-TrkA cells. a. scatchard plot analysis of high affinity ^{125}I[NGF] binding data; b. displacement of specific ^{125}I[NGF] binding by unlabeled NGF or by mimetic 1-ss-F*; c. mAb 5C3 competes for binding of FITC-tagged mimetic 1-ss-F*..... 39

Figure 2.6. Induction of putative TrkA dimers. After exposure to the indicated ligand (30 min at 4°C), NIH-TrkA cells were chemically cross-linked with DSS. After washing, they were detergent solubilized and samples (20 mg protein/lane) were studied by western blotting with a highly specific anti-TrkA mAb 5C3..... 41

Figure 2.7. CD Mimetic 1-ss-TEG antagonizes NGF and TrkA in cell survival assays. Mimetics were tested at 20 μM for antagonism of a. 0.2 nM NGF for TrkA cells; b. 0.1 nM NT-3 for TrkC cells; or c. accelerated cell death in SFM. Antagonism of d. NGF; or e. TrkA baseline activity were dose-dependent..... 44

Figure 2.8. Mimetic 1-ss-TEG antagonizes baseline (ligand-independent) and ligand-dependent TrkA activity. NIH-TrkA cells were cultured in SFM alone, and were then supplemented with the indicated ligand for 2 hours or for 20 min at 37°C in a cell incubator. After detergent solubilization the phosphotyrosine levels of the samples were studied by western blotting with anti-pTyr mAb 4G10. Total TrkA loading was verified on the same membranes with a highly specific anti-TrkA mAb 5C3..... 47

Page

Figure 2.9. Molecular modeling illustration of NGF:TrkA and 1-ss:TrkA interactions. a. ribbon diagram representation of TrkA dimer (green) bound to NGF dimer (cyan and light blue). The capped stick amino acids of the receptor indicate where monovalent D3 binds, near the NGF hot spot; b. the red box representing the putative binding pocket of compound 1-ss, used for docking; c. the surface of TrkA-D5 (green) with 1-ss shown in capped stick representation; d. a close-up of the circled area in c as representation of TrkA-D5 (green) complexed with 1-ss. The residues of TrkA-D5 forming hydrogen bonds with 1-ss (original atom color) were represented by capped stick models. The hydrogen bonds between 1-ss and TrkA-D5 are shown by dashed lines..... 49

Figure 3.1. a. The key distance of C^b-separations of the $i + 1$ to $i + 2$ residues of a type I β-turn and of the monovalent turn mimics DKP A and B featured here; b. an overlay comparison of the two types of 2,5-DKP mimics onto a type I β-turn. 55

Figure 3.2. Purities of the library of compounds 9 where the UV detection method was set at 254 nm. a. crude purity; and b. after purification. The term 'cap' is used for morpholine. 63

Figure 3.3. Structures of compounds 9 that bind to TrkC from the preliminary FACScan assay. 65

Figure 3.4. FACscan assay data and the graph presentation of the four "hits"..... 66

Figure 3.5. Structures of compounds 9 that block the trophic activity of NT-3 in TrkC cells in cell survival assay (9gg is shown in Figure 3.3)..... 67

Figure 3.6. Cell survival assay data and the graph presentation of the three "hits"..... 68

Figure 4.1. Three heterocyclic ring systems as Xaa-Gly dipeptide mimetics..... 71

Figure 4.2. Examples of different azole type compounds as peptidomimetics. a. non-peptidic tricyclic oxazole-pyrrole-piperazine scafold as α-helix mimetic; b. triazole ring-based new β-turn mimic.. 71

Figure 4.3. Examples of bioactive 1,3,4-oxadiazoles..... 72

Figure 4.4. a. The key distance of C^b-separations of the $i + 1$ to $i + 2$ residues of a type I β-turn and of the monovalent turn mimics 12; b. an overlay of the 1,3,4-oxadiazole mimic onto a type I β-turn. 73

Figure 4.5. a. Reaction of monomer 11i with DTAF (1:1) and the HPLC trace (UV

at 254 nm) of its crude purity; b. reaction of monomer 11s with DTAF (1:1) and the HPLC trace (UV at 254 nm) of its crude purity..... 81

Figure 5.1. Schematic view of Y-shaped IgG... 88

Figure 5.2. Structure of 22/8 and docking of 22/8 on IgG/protein A binding pocket... . 90

Figure 5.3. a. Fragment B of protein A complexed with IgG; b. an expanded view of the same interaction from a different perspective; and, c the protein A helix-loop-helix region with the residues that make contact with protein A highlighted.... 91

Figure 5.4. Cyclic peptidomimetics 19a-i of protein A... 93

Figure 5.5. a. Low energy conformer of compound 19a from quenched molecular dynamics (QMD) simulations; b. protein A fragment in complex with IgG; c. the best docking conformation of compound 19c on protein A backbone..... 94

Figure 5.6. ViewerPro plot of interaction details between domain III of protein G and the C_H1 domain of Fab. Hydrogen bonds between these two domains are marked using dashed lines. a. the overall interaction; b. the contact between the Fab and protein G β-sheets; and c. is the interaction between a β-strand of Fab (on the left) and the α-helix of domain III (on the right).... 99

Figure 5.7. Heterocyclic small molecules 22 and 23 as protein G mimetics... 101

Figure 5.8. Structure representation of active ligand fuctionalized dendrimers 25a-d on a DAB-Am-4 core... 106

LIST OF TABLES

Page

Table 2.1. Summary of the data of nine bivalent peptidimimetics 1a-i. 35

Table 3.1. Monovalent peptidomimetics 4c-d, 5c-d, 6a-b and 7a-b prepared via method a and b. 59

Table 4.1. Summary of 1,3,4-oxadiazole library 12a-t. 76

Table 5.1. Summary of cyclic peptidomimetics 19a-i. 97

Table 5.2. Summary of pyrrole and imidazole derivatized ligands 22 and 23..... 102

Table 5.3. SAR signals from sample injections. 105

Table 5.4. Summary of the active ligand-functionalized dendrimers 25a-d. 108

LIST OF SCHEMES

Page

Scheme 1.1. Synthesis of oxazole-based dipeptide mimetics as useful building blocks. 8

Scheme 1.2. Solid-phase synthesis of 1,3-azole-based peptidomimetics. 9

Scheme 1.3. A 'one-pot' synthesis of 1,2,4-oxadiazole-based peptidomimetics. 9

Scheme 1.4. 'One-pot' synthesis of highly substituted pyrazole library from 1,3 diketones. 10

Scheme 1.5. Combinatorial organic synthesis of 4-thiazolidinones and derived from amino acids 11

Scheme 1.6. Synthesis of *N*-terminal 2-iminohydantoins from di-/tripeptides and phenyl isothiocyanate. 12

Scheme 1.7. Synthesis of thiazolidine-based dipeptide mimetics from amino acid derivatives. 12

Scheme 1.8. Combinatorial synthesis of oxazoline and dihydrooxazine libraries from amino acid derivatives. 13

Scheme 1.9. Synthesis of γ-lactams by a tandem reductive amination/lactamization sequence. 14

Scheme 1.10. Combinatorial synthesis of piperazinediones and reduction to piperazines. 16

Scheme 1.11. Synthesis of 1,3,5-triazepan-2,6-dione library as dipeptide mimetics. 18

Scheme 2.1. Solid phase synthesis of the protected monovalent peptidomimetic A 31

Scheme 2.2. Overall scheme of preparation of the tagged bivalent peptidomimetics 1. 33

Page

Scheme 3.1. Two methods for preparing the monovalent mimics 4 and 5. Route a. featuring cyclo(Gly-Lys) and cyclo(Thr-Gly) monomers; Route b. featuring cyclo(Glu-Lys) and cyclo(Ile-Lys) monomers; c. linker attachment (for 4a-b and 5a-b, there is an additional hydrogenolysis step afterwards using H_2 and Pd-C in MeOH to remove side-chain protections to give 6b and 7b; and a Boc protection in CH_2Cl_2 to give 6a and 7a)..... 57

Scheme 3.2. Solution phase method for the preparation of the biotin-labeled library of bivalent compounds 9. 60

Scheme 4.1. General synthetic scheme of 1,3,4-oxadizole mimetics. a. synthesis of amino acid hydrazide fragments; b. synthesis of 1,3,4-oxadizole intermediates 11a-q; c. synthesis of monovalent oxadiazole mimics 12a-q. 74

Scheme 4.2. Synthesis of *N*,*N*-dimethyl serine methyl ester 10g and intermediates 11r-t. 76

Scheme 4.3. Preparation of bivalent mimetics 15. 79

Scheme 5.1. Synthesis of cyclic hexapeptide 19a. 96

Scheme 5.2. Synthetic scheme of pyrrole and imidazole derivatives. 102

Scheme 5.3. Synthesis of trivalent ligands 25a-d. a. hydrolysis of ester 22d, 22i, 22j and 22k; b. attachement of monovalent ligand 24a-d to a DAB-Am-4 core..... 107

CHAPTER I

INTRODUCTION

1.1 Heterocyclic Amide Surrogates of Dipeptide Mimetics

Many protein-protein interaction events are mediated through secondary structure units like helices, β-sheets and turns. To improve understanding of these molecular interactions, studies have been performed via mimicry of the crucial interacting sites. Small peptides and peptidomimetics are typically used to mimic or disrupt the structural environment at hot-spots.

Hetercycles are frequently found in natural and synthetic bioactive compounds. Heterocyclic building blocks can be useful as scaffolds for peptide backbone modifications. They import rigidity and enable the functionalized side-chain substituents to adopt relatively constrained conformations. Conceptually rigid analogs also surrender less entropy upon binding to their receptors. Many reports describe using heterocycles as successful peptide bond surrogates or protein recognition motif mimics to achieve ideal potency in biological assays.

Dipeptide mimetics can be accessed from conventional procedures and show various functionalities. They can also be used as protein secondary structure mimics and incorporated in active sequences for structural modification. Here, we mainly address the heterocyclic dipeptide mimetics.

This dissertation follows the style of the *Journal of Organic Chemistry*.

In general, we think the incorporation of amino acid side-chains, especially the ones that occur most frequently at hot-spots, is important for the mimic; it is also a direct means to introduce diverse functionalities to the heterocyclic system. In addition, the two "branches" should have some degrees of freedom for bond rotation; this allows the mimics to be somewhat flexible to be able to access close conformations to those of the parental peptides in binding protein targets. On the other hand, too much flexibility may cause the entropic cost of binding at a protein interface to be prohibitively high. A schematic view of the heterocyclic dipeptide mimetics is shown in Figure 1.1.

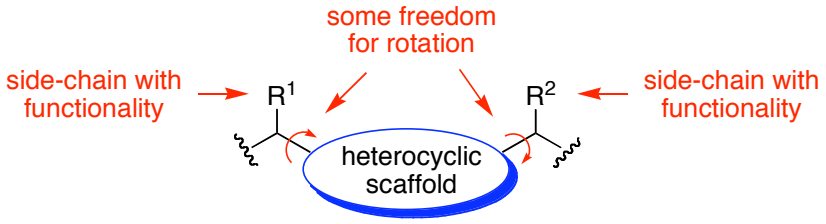
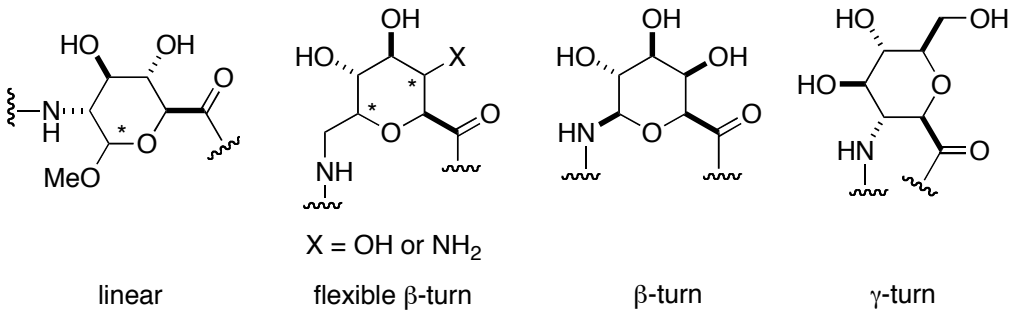


Figure 1.1. Schematic view of dipeptide mimetics based on a heterocyclic amide bond surrogate featuring two amino acid side-chains.

There are many types of dipeptide mimetics reported in the literature; they usually are rationally designed molecules and exhibit interesting characteristics, *e.g.* carbohydrate-based dieptide isosteres.[1] (Figure 1.2a) Sugar amino acids (SAAs) with naturally occurring rigid ring systems resemble conformationally restricted dipeptides, thus they are used as turn mimics and scaffolds to replace amide bonds in linear peptides.[2,3] Usually, specific functionalities are introduced at the sugar hydroxyls, or via backbone modifications.

a.



b.

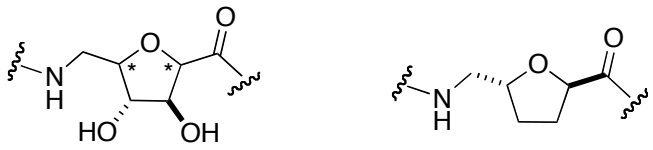


Figure 1.2. Various sugar amino acids as **a.** dipeptide isosteres;[1] and **b.** Gly-Gly replacement in Leu-enkephalin analogs.[4]

Bicyclic dipeptide mimetics are common peptidomimetic species. (Figure 1.3) They normally feature rigid backbones, which is crucial to maintain restricted conformations, and several substitution sites to access more molecular diversity. Not to mention many of them are also interesting targets for organic synthesis and biologically active compounds in medicinal chemistry.

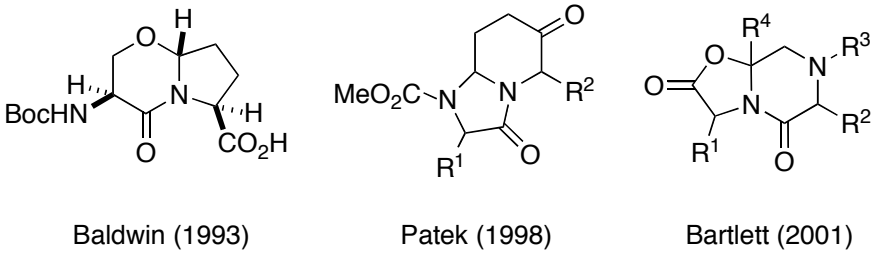
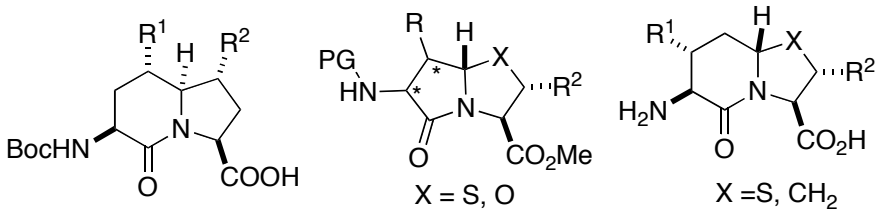


Figure 1.3. Various conformationally constrained bicyclic dipeptide mimetics.[5-12]



Lubell (2001) Hruby (2001, 2004) Hruby (2004, 2006)

Figure 1.3. continued.

Many dipeptide mimetics have been reported in the literature, however, only heterocyclic dipeptide mimetics that are pertinent to our research are described in detail in this chapter.

1.2 Five-membered Heterocyclic Dipeptide Mimetics

1.2.1 Azole-based Dipeptide Mimetics

Several types of five-membered heterocycles are used as amide bond surrogates of dipeptide mimetics, and many of them are based on azoles. Luthman and coworkers reported the synthesis of enantiopure heterocyclic Phe-Gly dipeptidomimetics containing 1,3,4-oxadiazole, 1,2,4-oxadiazole, and 1,2,4-triazole ring systems as building blocks.[13] (Figure 1.4) These pseudopeptides were used as Phe-Gly replacements in the biologically active peptides dermorphin (Tyr-D-Ala-Phe-Gly-Tyr-Pro-Ser-NH_2) and substance P (Arg-Pro-Lys-Pro-Gln-Gln-Phe-Phe-Gly-Leu-MetNH_2, SP). In the biological test of the μ- and δ-opioid receptor affinities of the dermorphin pseudopeptides, all mimetics except one type (1,3,4-oxadiazole, n = 0) displayed excellent affinities for the μ-receptor (IC_{50} = 12-31 nM) in the same range as dermorphin itself (IC_{50} = 6.2 nM); the tested pseudopeptides also retained their agonist activity at human μ-receptors. However, the SP pseudopeptides showed considerably lower affinities (IC_{50} > 1 μM) for the NK1 receptor than SP itself (IC_{50} = 1.5 nM), which indicated that the heterocyclic replacements prevent the pseudopeptides from adopting bioactive conformations.

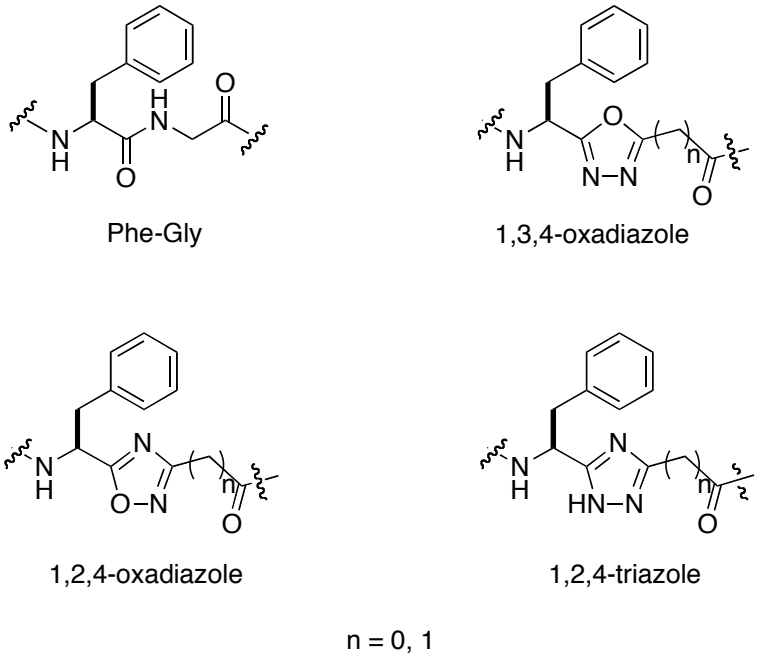


Figure 1.4. Three types of pseudopeptides as Phe-Gly replacements.

"Click" chemistry has made triazoles more accessible in recent years. They can be prepared from robust methods and have become versatile building blocks in organic synthesis. Triazoles have various applications, *e.g.* peptide amide bond surrogates.[14,15] Last year, Raines's group prepared several 1,4- and 1,5-disubstituted 1,2,3-triazoles as *cis*-peptide bond surrogates.[16] (Figure 1.5) The 1,4-disubstituted triazoles were prepared via Huisgen's 1,3-dipolar cycloaddition reaction, and the 1,5-disubstituted triazoles were prepared via Ru(II)-mediated 1,3-dipolar cycloaddition of various amino alkynes and azido acids. In the subsequent tests, they found that Xaa-1,5-triazole-Ala modules can serve as viable mimics of Xaa-*cis*-Pro segments in a protein, and the 1,5-regioisomers are superior to the 1,4-regioisomers in mimicking the *cis*-prolyl bond.

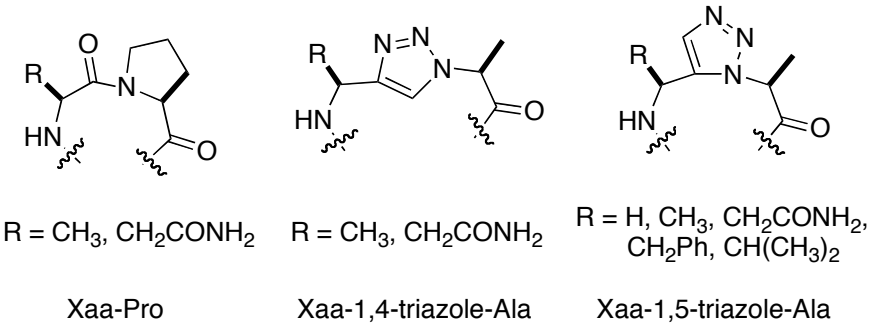


Figure 1.5. Xaa-1,5-triazole-Ala and Xaa-1,4-triazole-Ala modules as mimics of Xaa-*cis*-Pro segments.

To improve the properties of a pseudotetrapeptide FR-139317, which is a potent and highly selective antagonist of the endothelin-A (ET_A) receptor, von Geldern and his coworkers replaced Leu-Trp and Trp-X of the amide bond framework of FR-139317 with two different heterocyclic surrogates, respectively.[17,18] (Figure 1.6) The replacements generated two novel series of azole-based ET_A selective antagonists, but show a substantially different structure-activity profile from the peptidic series, particularly with regard to the side chain group incorporated into the heterocycle. The introduction of a heterocycle itself also has profound effects on the activity of the compounds.

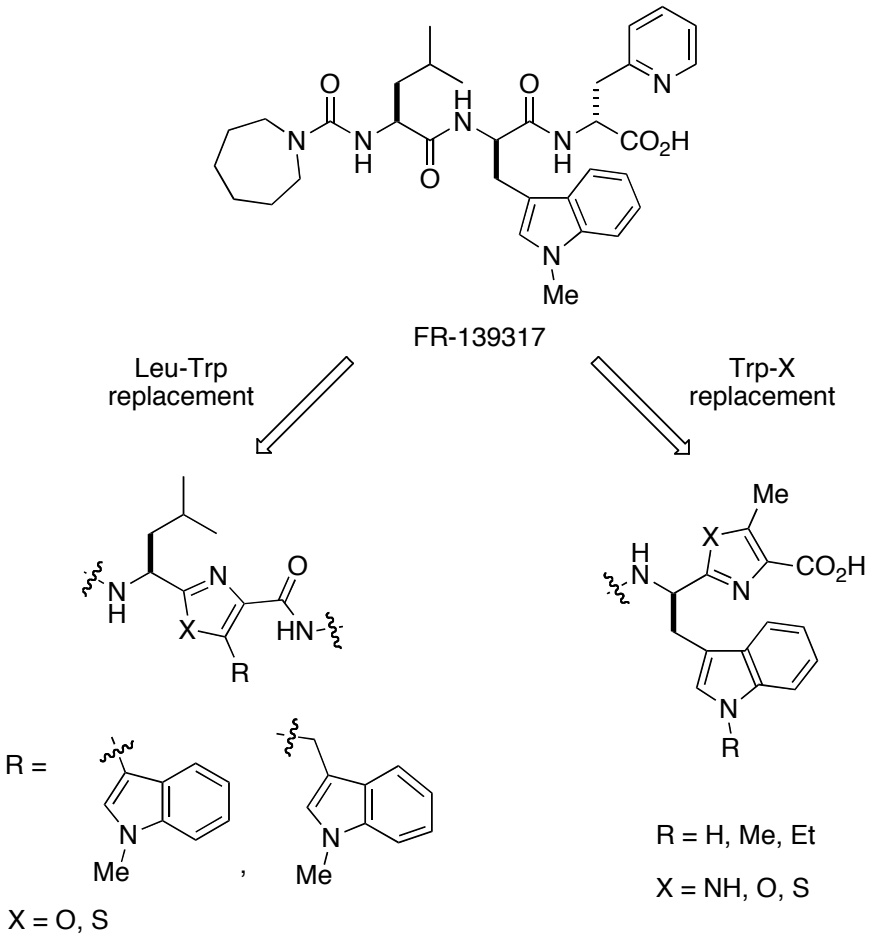
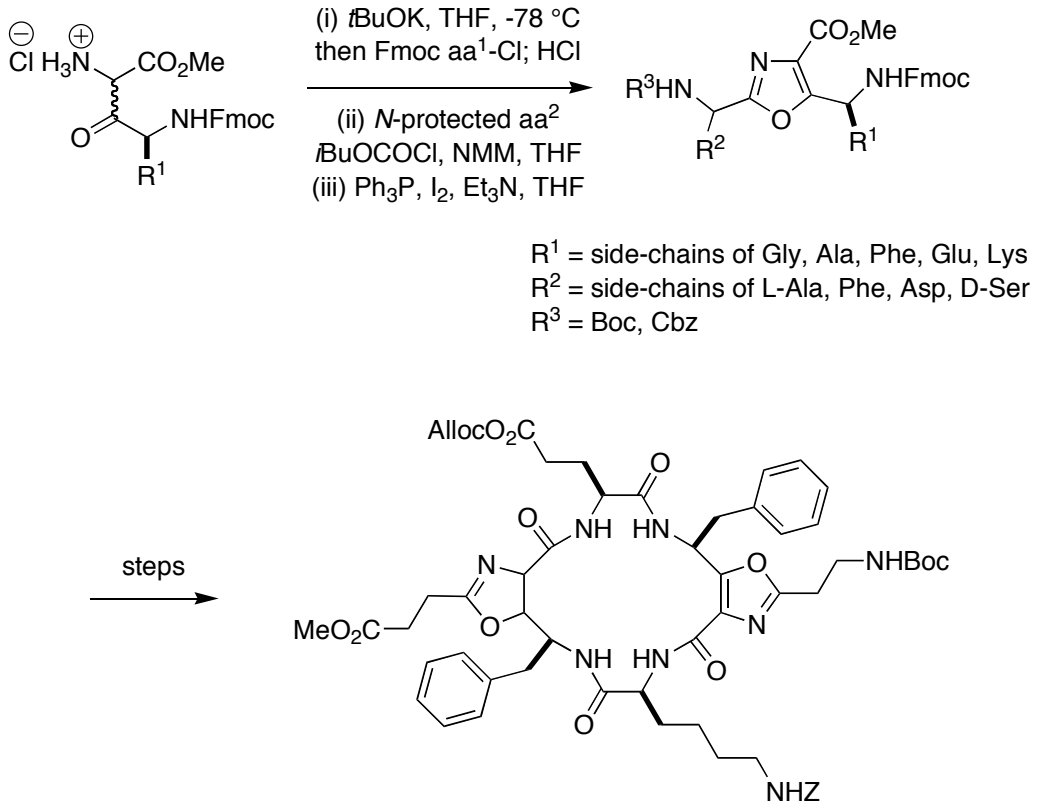


Figure 1.6. Azole-based Leu-Trp and Trp-X surrogates in potential endothelin antagonists.

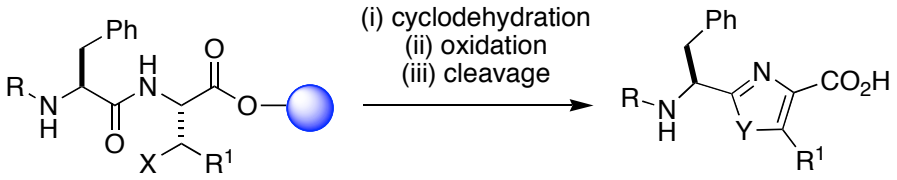
Kessler's group reported the synthesis of a new family of oxazole-based peptidomimetics as useful building blocks for the synthesis of orthogonally protected macrocyclic scaffolds.[19] (Scheme 1.1) These highly functionalized oxazoles were prepared from α-amino β-keto esters by acylation of the anion derived from *N*-(diphenylmethylene)-glycine methyl ester, using Fmoc-protected amino acid chlorides as acylating agents. Subsequent coupling with another amino acid with mixed anhydride activation afforded the corresponding intermediate, and the following cyclization yielded the desired oxadizole rings. These building blocks were used for preparing macrocycles containing Lys and Glu residues, which are presented as orthogonally protected scaffolds for supramolecular chemistry.

Scheme 1.1. Synthesis of oxazole-based dipeptide mimetics as useful building blocks.



The same group also reported an efficient two-step solid-phase synthesis of 1,3-azole-based peptides and peptidomimetics.[20] (Scheme 1.2) A series of 1,3-oxazole-, thiazole-, and imidazole-containing peptides were prepared from dipeptides that were composed of *C*-terminal threonine, serine, cysteine, or diaminopropionic acid via different cyclodehydration procedures followed or preceded by oxidation. The methods are compatible with Fmoc solid-phase peptide synthesis (Fmoc-SPPS) conditions. The reported methodology provides an efficient and flexible means to synthesize azole-based derivatives and allows the solid-phase synthesis of natural product libraries, small molecule combinatorial libraries, and 1,3-azole-based peptidomimetics. There was no report of the biological activities of these synthetic molecules.

Scheme 1.2. Solid-phase synthesis of 1,3-azole-based peptidomimetics.

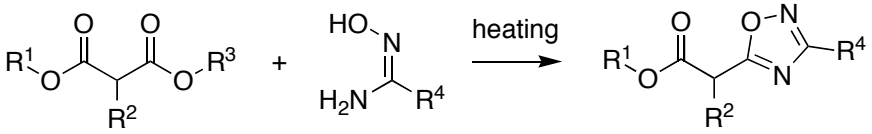


R^1 = H, Me
R = Boc, Fmoc, Z or Alloc
X = OH, STrt, NHTs

R^1 = H, Me
R = H, Fmoc, Z or Alloc
Y = O, S, NH

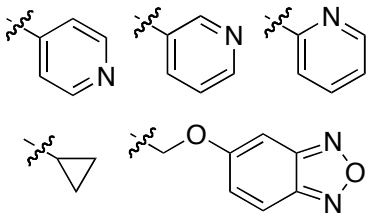
Du *et al* reported a convenient 'one-pot' synthesis of a series of oxadiazole-based peptidomimetics.[21] (Scheme 1.3) The α-1,2,4-oxadiazolo esters were synthesized from malonic diesters and amidoximes under solvent-free conditions. It is likely that the reaction goes through a ketene intermediate generated from the malonic diester by elimination of a molecule of alcohol. This strategy is both time and cost efficient compared to previously described methods. No biological activities of these molecules was reported.

Scheme 1.3. A 'one-pot' synthesis of 1,2,4-oxadiazole-based peptidomimetics.



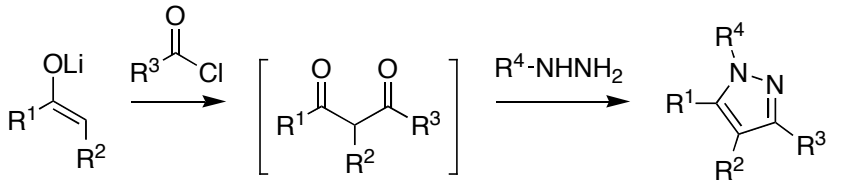
R^1 = CH_3, CH_2CH_3, $PhCH_2$
R^2 = H_2, $CH_2CH(CH_3)_2$, $PhCH_2$, Ph, C_5H_9, $(CH_3)_2$
R^3 = CH_3, CH_2CH_2, $C(CH_3)_3$

R^4 = CH_3, $CO_2CH_2CH_3$,



Heller and coworker prepared a series of highly substituted pyrazoles from 1,3-diketones via a rapid and general 'one-pot' synthetic methodology.[22] (Scheme 1.4) The 1,3-diketone building blocks were synthesized directly from ketones and acid chlorides, and were then converted *in situ* into pyrazole products upon the addition of hydrazine. This method is fast, general, and chemoselective. Moreover, it allows for the synthesis of some previously inaccessible pyrazoles and several synthetically demanding pyrazole-containing fused ring systems. No biological assay result is reported in their paper.

Scheme 1.4. 'One-pot' synthesis of highly substituted pyrazole library from 1,3-diketones.



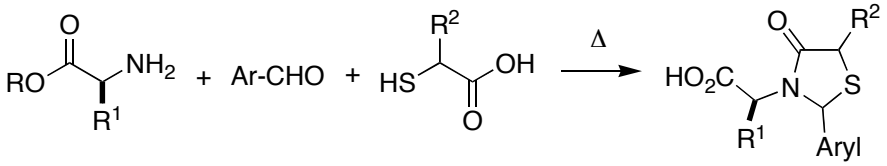
R^1 = Aryl
R^2 = H, Ph, *n*-Pr
R^3 = Aryl, alkyl
R^4 = H, Me, Ph

1.2.2 Other Five-membered Heterocycle-based Dipeptide Mimetics

Holmes *et al* reported the combinatorial synthesis strategies of 4-thiazolidinones and 4-metathiazanones derived from amino acids in both solution and solid-phase.[23] (Scheme 1.5) In solution, 4-thiazolidinones could be obtained from the 'one-pot', three-component condensation of amino acid esters, an aromatic aldehyde, and an α-mercapto carboxylic acid in moderate to high yields. The solid-phase approach to the five and six-membered heterocycles started from acylation of standard peptide synthesis resins with an Fmoc-protected amino acid, which was followed by Fmoc-removal and condensation with aldehydes and an α- or β-mercapto carboxylic acid. Cleavage from the support

using trifluoroacetic acid gave high yields and high purities of the liberated 4-thiazolidinones and lower yields of 4-metathiazanones. They did not report the biological properties of these compounds.

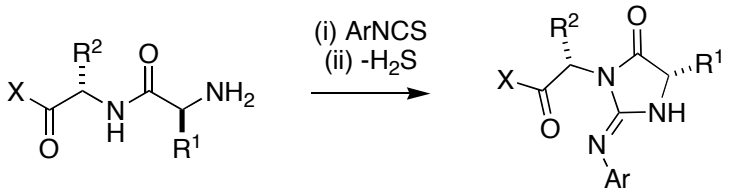
Scheme 1.5. Combinatorial organic synthesis of 4-thiazolidinones and derived from amino acids.



R^1 = side-chains of Gly, Ala, Phe, Val, β-Ala
R^2 = H, Me
R = alkyl ester or polymer support

Batey and coworker prepared of a series of *N*-terminal 2-iminohydantoin-based peptide heterocycle conjugates via a diverted Edman degradation protocol.[24] (Scheme 1.6) This modified procedure provides an effective route to introduce a heterocycle at the *N*-terminus of an α-amino acid amide or peptide. The iminohydantoins are generated from the reaction of a peptide with an isothiocyanate, followed by dehydrothiolative trapping of the intermediate thiourea, and intramolecular cyclization of the weakly nucleophilic adjacent amide nitrogen. A solution-phase parallel synthesis of the peptidomimetics and a solid-phase synthesis of dipeptide- and tripeptide-derived iminohydantoins were developed. They did not report the biological activities of these molecules.

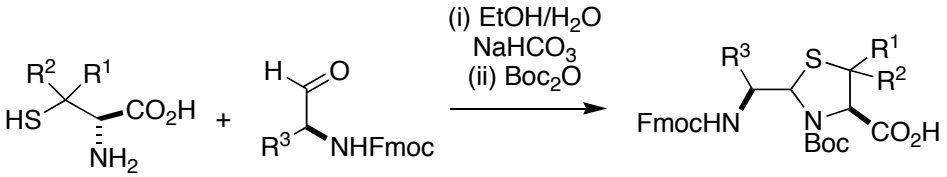
Scheme 1.6. Synthesis of *N*-terminal 2-iminohydantoins from di-/tripeptides and phenyl isothiocyanate.



R^1 = side-chains of Leu, Ala, Phe
R^2 = side-chains of Phe, Leu, Ala
X = O*t*Bu, Phe-O*t*Bu, Leu-O*t*Bu

Grieco *et al* reported the design and synthesis of small libraries of peptidomimetics based on a thiazolidine moiety.[25] (Scheme 1.7) In this work, two series of conformationally constrained building blocks starting from amino acids were prepared via a short and general strategy. Computer methods were used to evaluate the intrinsic tendency of the dipeptide mimetic scaffold to induce β-turn formation; the results indicate that the diastereomers with the (*R*) configuration at the C-2 position have high propensity to adopt the type II β-turn conformation.

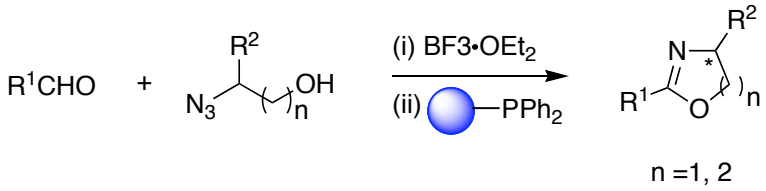
Scheme 1.7. Synthesis of thiazolidine-based dipeptide mimetics from amino acid derivatives.



R^1 = R^2 = H, CH_3
R^3 = side-chains of Gly, Ala, Asp, Lys, Phe, Tyr, Trp, His, etc

Last year, Aubé's group reported the one-step synthesis of a 60-member library of various disubstituted oxazolines and di- and trisubstituted dihydrooxazines.[26] (Scheme 1.8) The reactions of 1,2- and 1,3-hydroxyalkyl azides and aldehydes in the presence of Lewis acid generate oxazolines and dihydrooxazines in one step, respectively. Parallel synthesis was adapted using a polymer-bound phosphine to scavenge excess hydroxyalkyl azide. The activities of the library members were being evaluated against various biological screens.

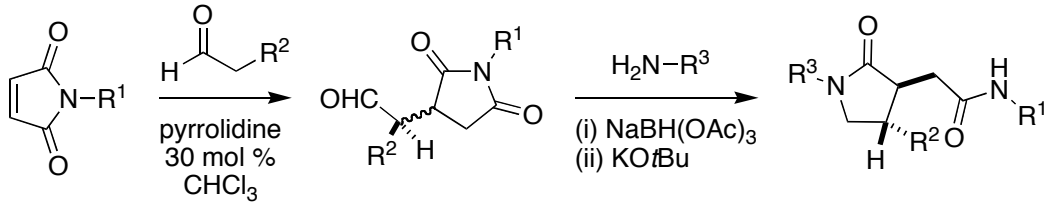
Scheme 1.8. Combinatorial synthesis of oxazoline and dihydrooxazine libraries from amino acid derivatives.



R^1 = Aryl
R^2 = side-chains of Phe, Val, Leu, Asp(OBn), Ph, etc

This year, Reiser *et al* reported an efficient synthesis of γ-lactams by a tandem reductive amination/lactamization sequence.[27] (Scheme 1.9) This three-component method allows for the synthesis of highly substituted γ-lactams from readily available maleimides, aldehydes, and amines. The reported synthetic strategy provides a straightforward route to the lactam products, and the parallel synthesis of a γ-lactam library further demonstrates its general utility. The biological activities of these molecules have not been reported yet.

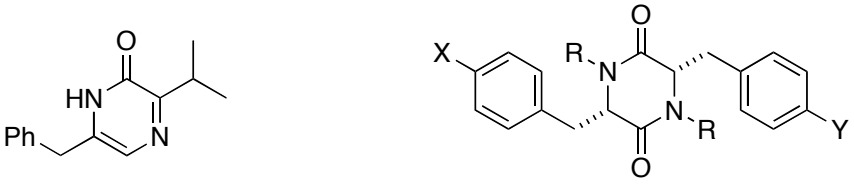
Scheme 1.9. Synthesis of γ-lactams by a tandem reductive amination/lactamization sequence.



R^1 = Ph, Bn, *p*-MeOBn
R^2 = Bn, Me, *n*-pentyl
R^3 = Aryl, alkyl

1.3 Six-membered Heterocyclic Dipeptide Mimetics

Diketopiperazines (DKPs) are useful templates for constructing dipeptide mimetics. Aubé's group reported the synthesis of a small library of 2,5-diketopiperazines as potential inhibitors of calpain.[28] (Figure 1.7) These constrained dipeptide analogs were designed based on previously reported calpain inhibitors, *e.g.* phevalin; they were synthesized starting from corresponding amino acid building blocks. A concise total synthesis of the structurally related natural product phevalin was also reported. However, despite the dipeptidyl aldehyde analog Boc-Val-Phe-H was a potent calpain inhibitor as expected, none of the library members, including previously reported phevalin, showed any significant inhibition against recombinant human calpain I in their assay system.



Phevalin
IC_{50} 1.3 μM

DKP (R = Me, X = Y= OH)
IC_{50} 0.8 μM

R = H, Me
X = H, OH
Y = H, OH, OBn

Figure 1.7. Dipeptide mimetics of calpain based on 2,5-diketopiperazines.

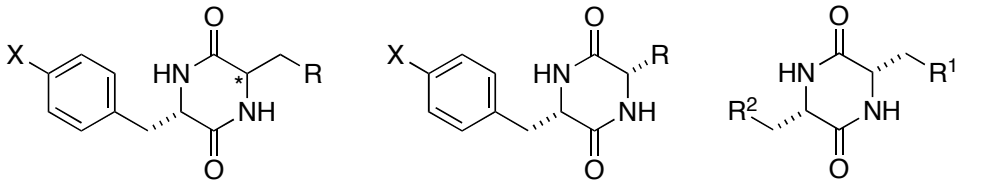
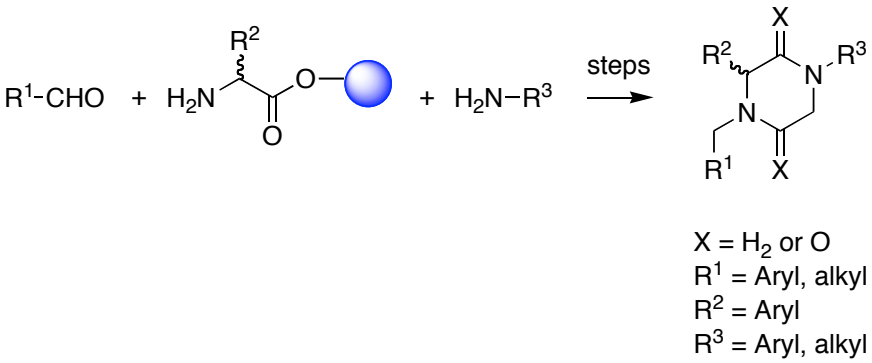


Figure 1.7. continued.

Goodfellow *et al* prepared a rationally designed library of variously substituted piperazines for the discovery of bradykinin antagonists and other G-protein-coupled receptor ligands on the basis of previous successful structures.[29] (Scheme 1.10) The synthetic strategy uses diverse nonnatural amino acids, on-resin-submonomer synthesis to provide diverse *N*-substituted structures, and the adaptation of simultaneous ring closure and resin cleavage drives the formation of highly hindered amide bonds. Thus a rationally directed library of ~ 2500 *N*,*N*'-disubstituted piperazines and piperazinediones was prepared and screened for ligand affinity on bradykinin, neurokinin, and opioid receptors. A number of lead structures, including a bradykinin antagonist lead CP-2458 (not shown), with good receptor selectivity and antagonist activity in human-cell assays were identified.

Scheme 1.10. Combinatorial synthesis of piperazinediones and reduction to piperazines.



1.4 Larger (>6) Heterocyclic Ring Systems as Dipeptide Mimetics

Burkholder and coworkers reported the synthesis of a series of 6-amino-5-oxo-7-phenyl-1,4-oxazepines as conformationally constrained *gauche* (-) Phe-X dipeptide mimetics, starting from β-phenylserine by a versatile route.[30] (Figure 1.8) To elucidate the solution conformation of the protected mimetics, NMR studies were conducted and showed that the 7-membered ring was in a chair conformation with the phthalimide moiety and the phenyl ring pseudo-equatorial, consistent with the *gauche* (-) conformation (dihedral angle X_1= -60°). These peptidomimetics may be useful probes for the investigation of conformational preferences of enzyme substrates and receptor ligands. MDL 100,192, as a *gauche* (-) Phe-Leu constrained mimetic, was reported to be a poor inhibitor of atriopeptidase (neutral endopeptidase 24.11, NEP) and angiotensin converting enzyme (ACE). The K_i for MDL 100,192 for NEP was 330 nM and the K_i for ACE was 50 μM.

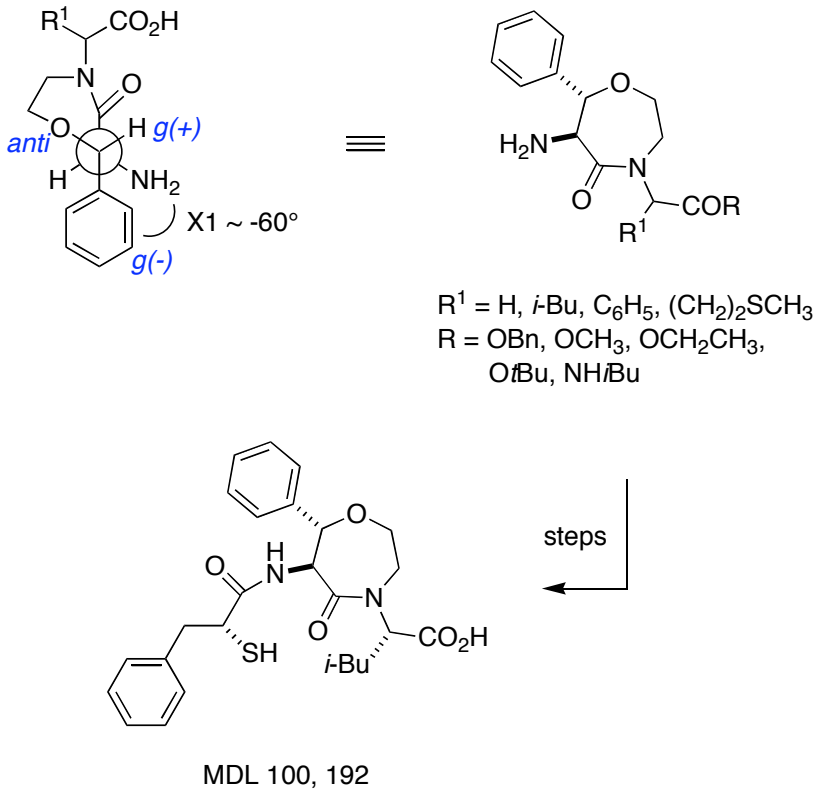
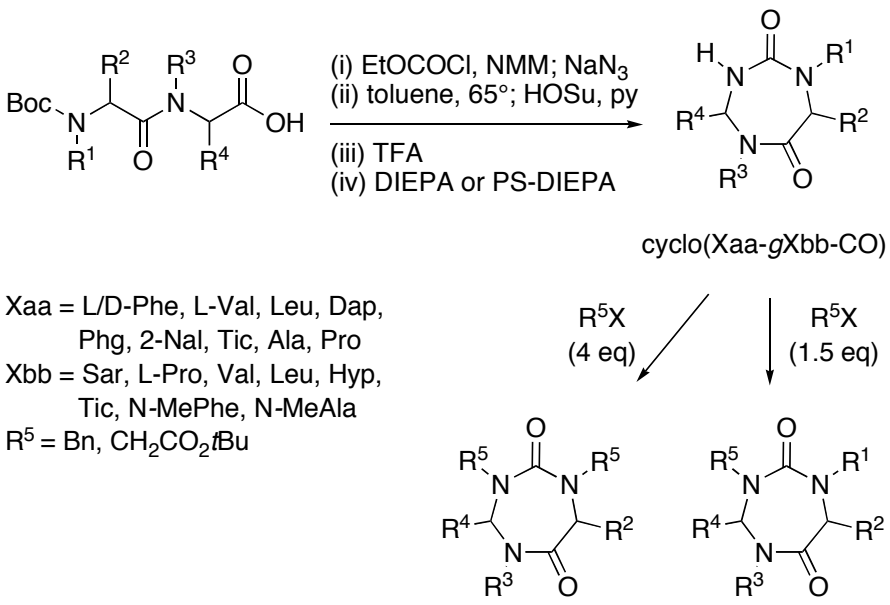


Figure 1.8. 1,4-Oxazepines as conformationally constrained *gauche* (-) Phe-X dipeptide mimetics.

Guichard *et al* reported the the design and synthesis of a structurally diverse small pilot library of 1,3,5-triazepan-2,6-diones as conformationally constrained dipeptide mimetics for identification of malaria liver stage inhibitors.[31] (Scheme1.11) The densely functionalized ring skeleton was readily accessible in only four steps from a variety of simple linear dipeptide precursors. A general polymer-assisted parallel solution-phase synthesis approach compatible with library production was developed. Conformational analyses by NMR spectroscopy and X-ray diffraction showed that the ring exhibits a characteristic folded conformation. Preliminary biological screening of a portion of the library revealed two molecules active against the malaria liver stages, which do not show any toxicity on mouse hepatocytes, and also suggested strong potential functionality of this dipeptide-derived scaffold in biological pathways.

Scheme 1.11. Synthesis of 1,3,5-triazepan-2,6-dione library as dipeptide mimetics.



In an early report by Olson *et al*, the design and synthesis of several protein β-turn mimetics based on a nine-membered-ring lactam framework was described.[32] (Figure 1.9) The synthesis of model di- and tetrapeptide mimetics started from 1,5-cyclooctadiene derivatives. In the model dipeptide mimetic, the amide linkage is *trans* based on NMR and X-ray analyses, and functional groups at positions adjacent to the lactam amide bond correspond closely to the side-chain positions of residues $i + 1$ and $i + 2$ of classical type II' β-turns. There is no biological data reported in this publication.

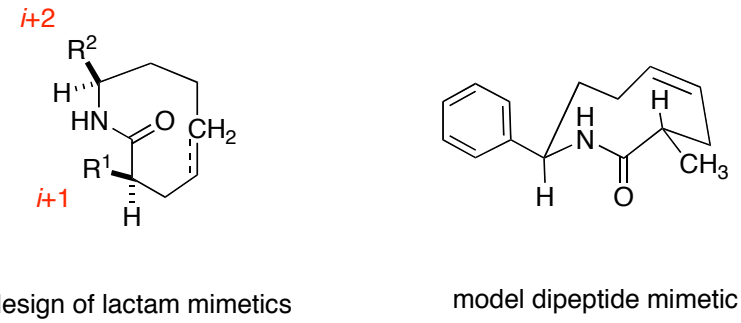


Figure 1.9. Nine-membered-ring lactam-based dipeptide mimetic as β-turn analog.

1.5 Other Heterocylic Dipeptide Mimetics

There are some examples of heterocyclic dipeptide mimetics that are not covered in previous sections mainly because of lack of space. These are also rationally designed molecules. Some are secondary structure mimics, and some exhibit interesting activities in various biological systems. Several examples are shown in Figure 1.10.

a.

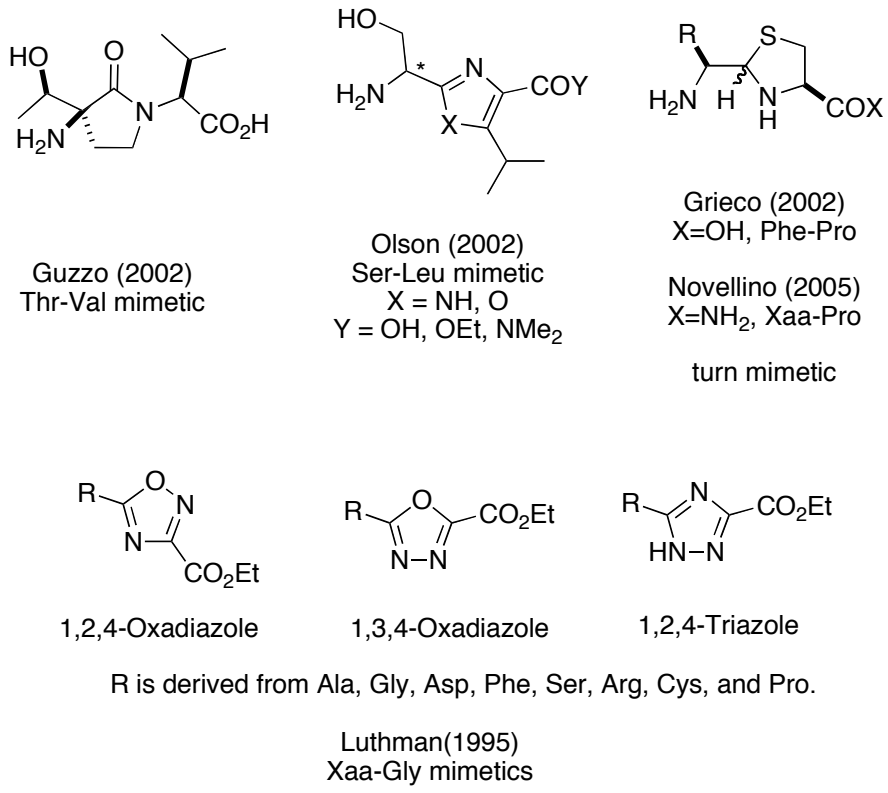
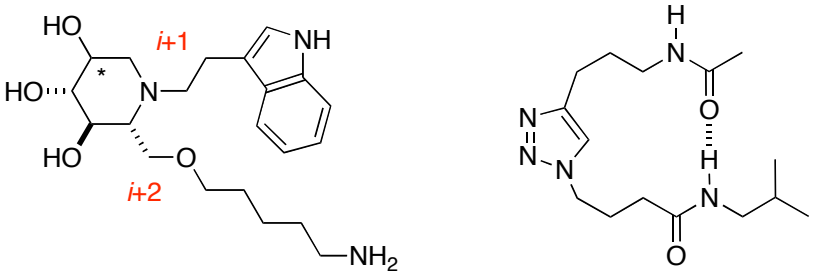


Figure 1.10. Various heterocyclic dipeptide mimetics. **a.** dipeptide mimetics mimicking particular dipeptide sequences;[33-37] **b.** dipeptide mimetics mimicking secondary structures;[38,39] **c.** conformationlly restricted dipeptide mimetics.[40-42]

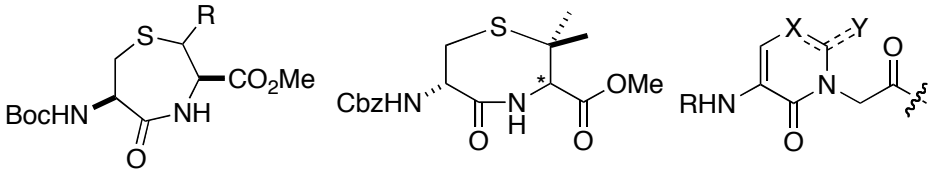
b.



Murphy (2008)
Trp-Lys mimetic
β-turn mimetic

Guan (2006)
β–turn mimetic

c.



Botta (1999)
R = α-Me, β-Me

Goodman (1996)

Tamura (1997)
X =CH, NCH_3, N
Y = H, CH_3, O
R = SO_2Bn

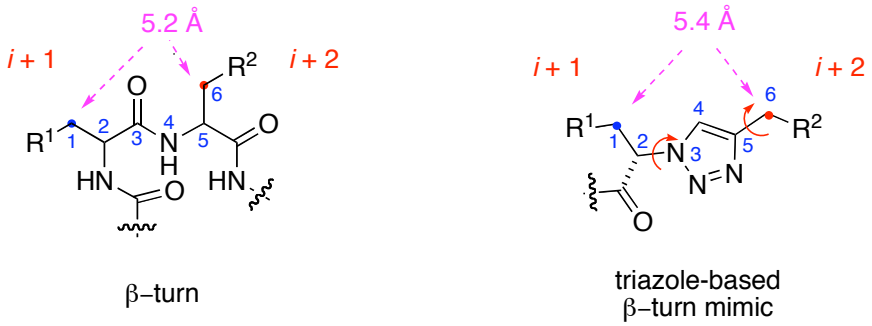
Figure 1.10. continued.

1.6 Monovalent Heterocyclic Dipeptide Mimetics Designed in Our Group

The design of our previously reported triazole-based β-turn mimics (Figure 1.11a) was based on two major considerations.[43] First, studies of protein complexes crystallographically have shown that main-chain carbonyl groups are much less (~ 11 %) involved in protein-protein interface regions than side-chains (~ 80 %). Thus we believed it is far more important to represent amino acid side-chains in the mimics, than main-chain amides. This is consistent with designs by some other groups, *e.g.* Hirschmann and Smith. Second, syntheses of the mimics should also incorporate the side-chains that occur most frequently at hot-spots, *i.e.* those of Trp, Arg, Tyr, Lys, Glu, Ser, Asn, Leu; syntheses of β-turn mimics that do not allow this are unsatisfactory for

medicinal chemistry. Actually, it is quite hard to devise syntheses of mimics that do satisfy this criterion because of the diversity of amino acid side-chain functionalities involved.

a.



b.

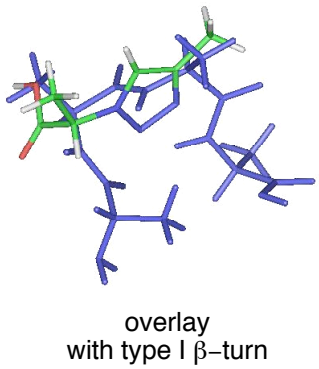


Figure 1.11. a. The key distance of C^{β}-separations of the $i + 1$ to $i + 2$ residues of a type 1 β-turn and of the triazole-based turn mimics featured here; **b.** an overlay of the mimics onto a type 1 β-turn.

Conformational and structural issues are also factors in our design considerations for these turn mimics. As shown in Figure 1.11a, the bond vectors 1-2, 2-3, 3-4, 4-5, and 5-6 populate relatively well-defined regions of space, whereas rotation about the 1-2 and 5-6 bonds in β-turns are relatively free. We also reasoned that good mimics should be able to access conformations with $C^1 - C^6$ separations that are close to those of β-turns.

Type-1 β-turn is known as the most common type among several different types of β-turn conformations. However, most β-turns probably are somewhat flexible in solution, thus good mimics should reflect that characteristic without being so flexible that the entropic cost of binding at a protein interface is prohibitively high. Although the conformation shown in Figure 1.11b is not the low energy conformation, which indicates the turn mimic does not always reside in it, this conformational state is readily accessible at room temperature since the enthalpic barrier to rotation about C^2-N^3 dihedral is low (not shown). Computer-aided calculations showed that the lowest energy conformer that was identified has a C^2-N^3 dihedral angle of 50° which is 140° away from the targeted turn conformation. However, the second lowest energy conformer was only 0.85 kcal•mol^{-1} higher in energy and had a C^2-N^3 dihedral angle of –57°, which matched the target turn conformation much closer (deviation = 33°). The calculations also showed that the molecule only need to surmount an enthalpic barrier of 0.99 kcal•mol^{-1} to reach the target turn conformation; which can be readily overcome at room temperature. Overall, the turn conformation of our peptidomimetics is energetically accessible, and it represents a low energy conformation that is only slightly less stable than the minima. The energy needed to reach this conformation is available at ambient temperatures, and the slight increase of energy over the lowest minimum may be easily compensated for via induced fit binding to a protein surface.

The designated mimics in Figure 1.11a satisfy all the considerations outlined above. The separations between C^β atoms that correspond to the $i + 1$ and $i + 2$ amino acid side-chains (*i.e.* C^1 and C^6) in type-1 β-turns are close to that in the mimics in an energetically accessible conformation. β-Turns have two bonds that are essentially freely rotating (1-2 and 5-6). The core of triazole-based molecules has three rotatable (1-2, 2-3, and 5-6), and only one of these affects the C^1-to-C^6 distance (the 2-3 rotation). Consequently, the peptidomimetics can access appropriate conformations to mimic the C^1-to-C^6 spacing in β-turns simply via rotation around the 2-3 bond (Figures 1.11a and b). In addition, the polar triazole part of the mimics introduces electrostatic polarity that can also resemble turn conformations (not shown). Thus, we think the design is satisfactory by the match of the mimics to the target secondary structure.

Later, our group devised the synthesis of another series of similar β-turn peptidomimetics (not published). These triazole-based molecules were also prepared via copper-mediated 1,3-dipolar cycloaddition reactions, but using only various natural amino acids derived alkynes and azido acids as building blocks. Their biological activities are still under investigation. We are currently working on the design and synthesis of other β-turn peptidomimetics based on different types of heterocycles.

1.7 Conclusion

To improve understanding of various protein-protein interactions, studies have been performed via mimicry of the crucial interacting sites. Small peptides and peptidomimetics are typically used to mimic or disrupt the structural environment at hot-spots.

Hetercycles are frequently found in natural and synthetic bioactive compounds. Heterocycle-based dipeptide mimetics usually feature relatively constrained conformations and various functionalities. Many reports describe using heterocyclic building blocks, such as azoles, azolines, and diketopiperazines, as successful amide bond surrogates for protein backbone modifications to improve their biological properties. These molecules can also be used to resemble protein secondary structures.

In general, we think the incorporation of the amino acid side-chains that occur most frequently at hot-spots in the mimics is quite important, and good mimics should also be somewhat flexible to be able to access favorable conformations that are close to those of the parental peptides in binding protein targets.

CHAPTER II

DIMERIZATION OF A MONOVALENT SMALL MOLECULE LIGAND CHANGES A TrkA RECEPTOR AGONIST INTO AN ANTAGONIST

2.1 Introduction

TrkA is a member of the receptor tyrosine kinase (RTK) family, and the high affinity receptor for Nerve Growth Factor (NGF). RTKs in general, and TrkA specifically, are believed to act via receptor dimerization, where each receptor trans-activates the juxtaposed intracellular tyrosine kinase domain of its neighbor.[44] This hypothesis fits well with the fact that NGF is a dimeric ligand, so it seems logical that it can induce ligand-dependent TrkA receptor dimerization or stabilize the preformed receptor dimers.[45] However, postulates based on the idea that RTK dimerization is the exclusive requirement for agonistic activities are inconsistent with the observation that *monovalent* ligands of RTK can be agonistic. [46-48] Actually in the specific case of TrkA, it has been shown that dimers exist on the cell surface before NGF is introduced. This fact further suggests ligand binding may cause conformational changes for activation of the receptor intracellular domain. Indeed, monovalent ligands of TrkA can be agonists of this receptor,[49,50] indicating that pre-formed TrkA dimers can be activated by ligands that are non-dimeric. Actually, monvalent and monomeric agonists have been reported for other RTKs,[51-53] thus this phenomenon is not unique to TrkA. This evidence supports the notion that conformational changes in TrkA receptor dimers may occur. However, conformational changes in TrkA upon the binding of protein ligand, such as NGF, have not yet been demonstrated through structural analyses.[54]

If the correlation between RTK dimerization and agonistic activity is still in question for protein ligands, it is even less certain for small molecule ligands. Small molecule ligands and proteins normally have different kinetic and thermodynamic profiles for binding to RTKs. Synthetic compounds can bind in allosteric sites, and

especially bivalent small molecules could bind not only one, but also two receptor molecules at the same time.

Our laboratory has been involved in the design and synthesis of small molecules to mimic surface exposed β-turn regions of NGF (and other neurotrophins).[55-60] Specifically, we have targeted mimics of the NGF turn regions that seem to be involved in those "hot-spot" interactions between NGF and TrkA. Most of these small molecules are cyclic peptidomimetics with our so-called "ring fused C^{10} motif" as illustrated in Figure 2.1. They were made via SPPS and S_NAr macrocyclization reactions. Residues at R^1 and R^2 positions were based on the $i + 1$ and $i + 2$ residues of the turn regions of NGF.

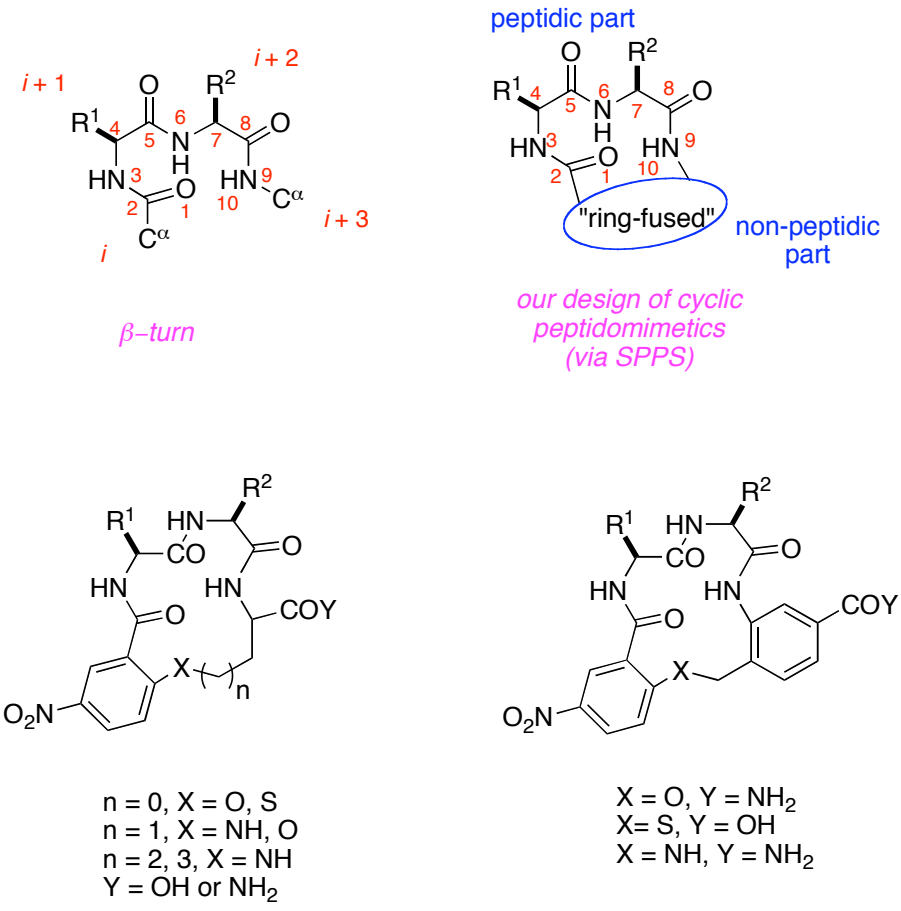


Figure 2.1. Some β-turn nomenclature and some cyclic peptidomimics with ring-fused C^{10} motif.

Among libraries of these monovalent cyclic peptidomimetics, we have identified several lead compounds based on extensive biological assays, such as **D3** and **MPT18** (Figure 2.2).[58,61] Interestingly, peptidomimetic **D3** was shown to have agonistic (or partial agonistic) properties for the NGF:TrkA interaction in biochemical (*e.g.* tyrosine phosphorylation), cellular (*e.g.* survival and neurite outgrowth),[50,62] and *in vivo* assays (Robin's maze test for gain of memory function in aged rats), which demonstrated that monovalent mimics on hot spot of NGF do not have to be antagonists.[63,64] Compound **D3** has also been shown to bind the TrkA receptor in a non-competitive manner respect to NGF.

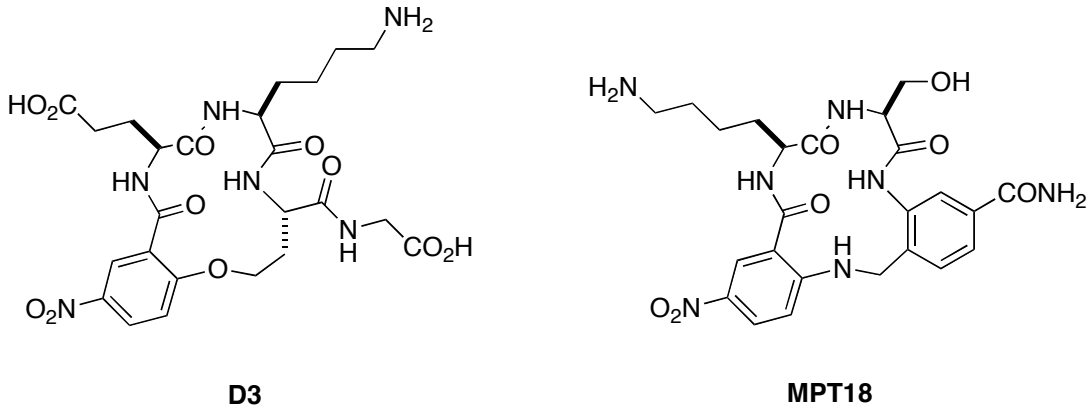


Figure 2.2. Previously identified lead compounds **D3** and **MPT18**.

Given the interesting biochemical and biological profile of peptidomimetic **D3** and the notion of changing a small molecule from a monovalent to bivalent form could generate more potent agonists for RTK ligands (actually more potent TrkA agonists have been generated by increasing ligand valency),[49,62,65] it seemed reasonable to explore "bivalent analogs" in which two close analogs of **D3** (which is **A** as shown in Figure 2.3a) are combined in one molecule.[66,67] However, what will be the ideal dimensions for the linker if they are to be combined? It is almost impossible to surely answer this question because the precise binding site of **D3** on TrkA is unknown, and also its bivalent analogs might bind differently in any case. Furthermore, the only fragments of the Trk

receptor structure that have been crystallographically characterized are not the regions that contact the turns that **D3** was designed to mimic. Consequently, our strategy was to generate a set of bivalent molecules **1** using a novel combinatorial approach. In this method, a close analog of **D3** (which is **A**) would be attached to a short **s**, a medium **m**, and a long linker **l**. They were then paired on a triazine scaffold via our previously published methodology, to give a small library of bivalent compounds representing all combinations of linkers of the different lengths, *ie* **1-ss**, **1-sm**, **1-sl**, **1-mm**, **1-ml**, and **1-ll**. (Figure 2.3b)

Two different "tags" were prepared in our lab and used on the triazine linker. A triethylene glycol (TEG) based one had no particular function except to increase the water solubility of the unlabeled mimics. However, fluorescein labeled compounds were critical because they facilitated screening via fluorescence activated cell sorting (FACS) using transfectant cells that overexpress TrkA; these experiments gave relative degrees of staining from which relative binding affinities could be inferred. Then we examined the TEG-tagged compounds as competing ligands in the same assay. Finally, the optimal binder was tested in: (i) competitive binding assays with radiolabeled NGF to obtain K_d values, and in: (ii) cell survival assays to test agonism or antagonistic effects for the NGF:TrkA interaction. As a result, the data obtained was surprising in two respects. Firstly, the bivalent analog **1-ss**, which featured the shortest linker combination, had the highest affinity to TrkA, when it might have been anticipated that those with longer, more flexible linkers would be better. Second, unlike its parental monovalent peptidomimetic **D3**, **1-ss** was shown to be an antagonist for the NGF: TrkA interaction in cellular assays.

a.

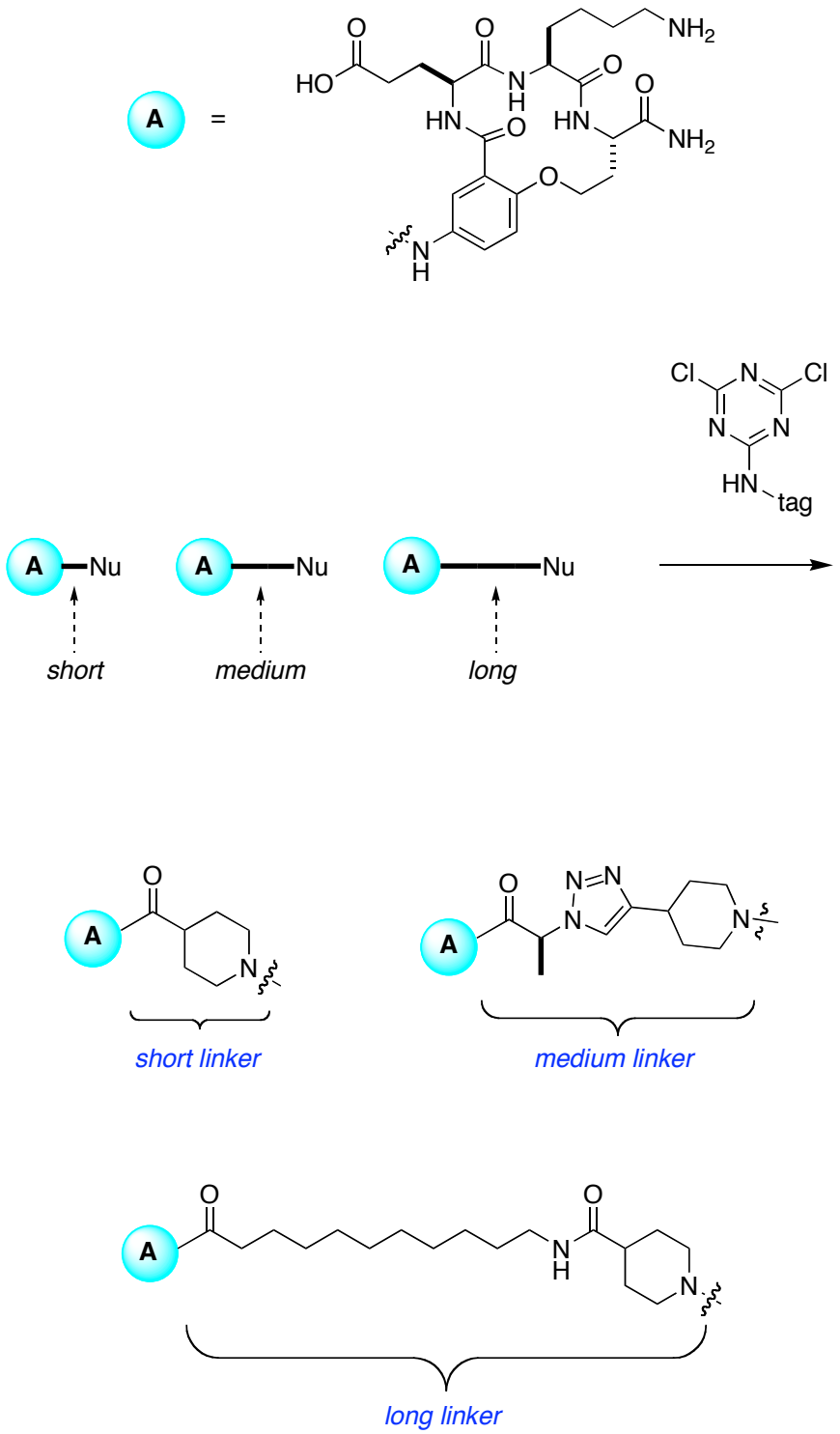
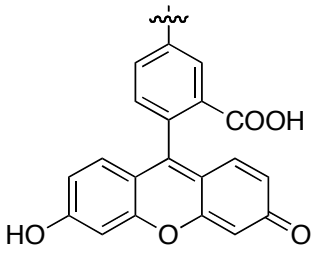
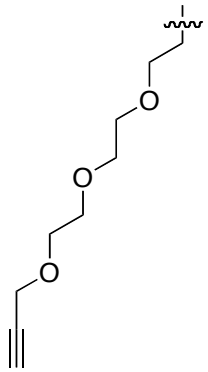


Figure 2.3. **a.** Generic scheme of the monovalent mimic **A** was attached to a short **s**, a medium **m**, and a long linker **l** bearing a nucleophile, then paired on a triazine scaffold to which either a FITC or a TEG tag was incorporated; **b.** structure of the bivalent compounds **1-ss**, **1-sm**, **1-mm**, **1-sl**, **1-ml**, and **1-ll**.

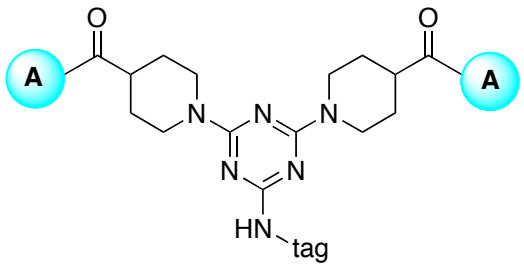


Fluorescein tag (F*)

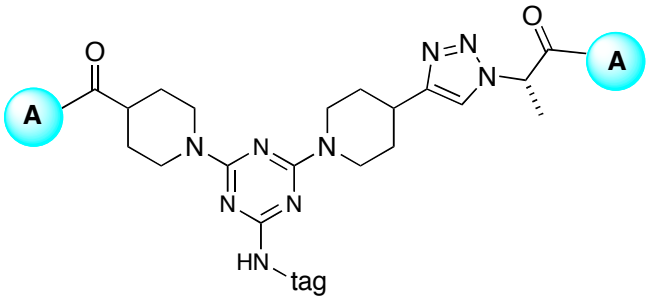


TEG tag

b.

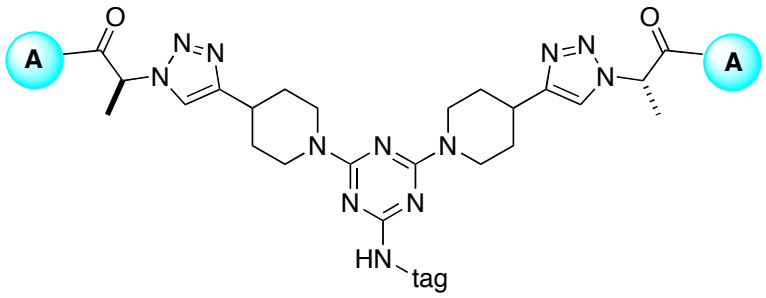


1-ss

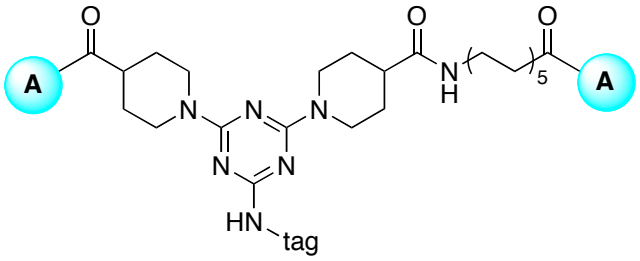


1-sm

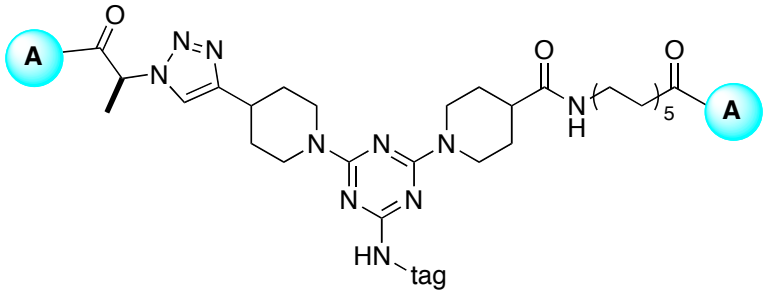
Figure 2.3. continued.



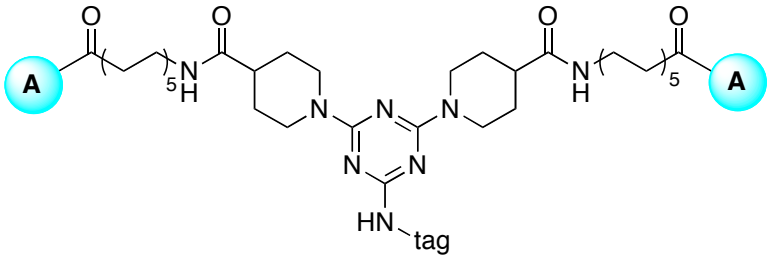
1-mm



1-sl



1-ml



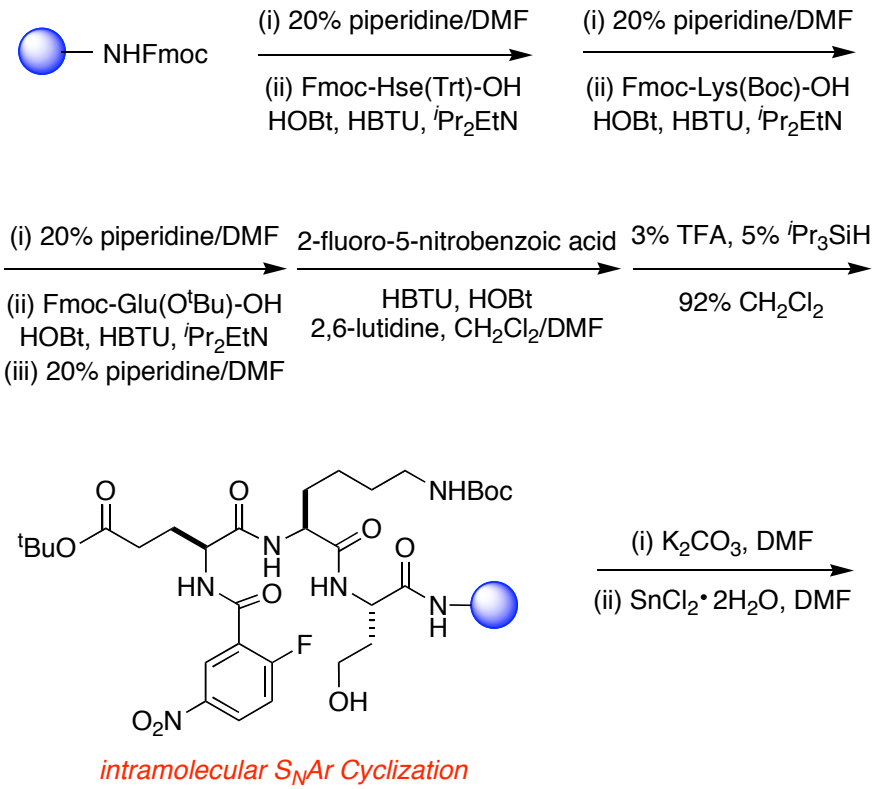
1-ll

Figure 2.3. continued.

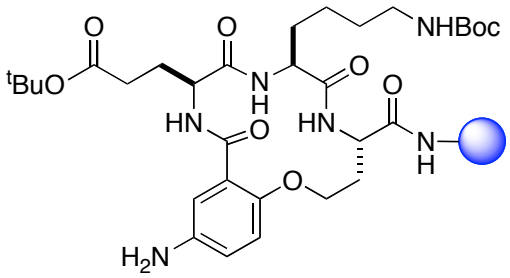
2.2 Solid Phase Synthesis of the Monovalent Cyclic Peptidomimetic A

Scheme 2.1 outlines the preparation of **A**, as a close analog of cyclic monovalent compound **D3**. The synthesis represents a slight modification of the literature procedure from our group.[68] FMOC-Hse-(Trt)-OH, FMOC-Lys(Boc)-OH and FMOC-Glu(O^{t}Bu)-OH were sequentially coupled onto the deprotected Rink Amide MBHA resin under typical solid phase coupling conditions. 2-Fluoro-5-nitrobenzoic acid moiety was introduced following Glu and the side-chain protecting group (Trt) of the homoserine was removed by treatment with acid to afford the linear peptide precursor. The macrocyclization step was carried out by treating the supported linear peptide with K_2CO_3 in DMF at room temperature for 1 day. The nitro group of the cyclic compound was reduced to an amino group using $SnCl_2$; this was used as a handle for further linker attachment.

Scheme 2.1. Solid phase synthesis of the protected monovalent peptidomimetic **A**.



Scheme 2.1. continued.



2.3 Synthesis of the Labeled Bivalent Peptidimimetics 1 with s/m/l Linkers

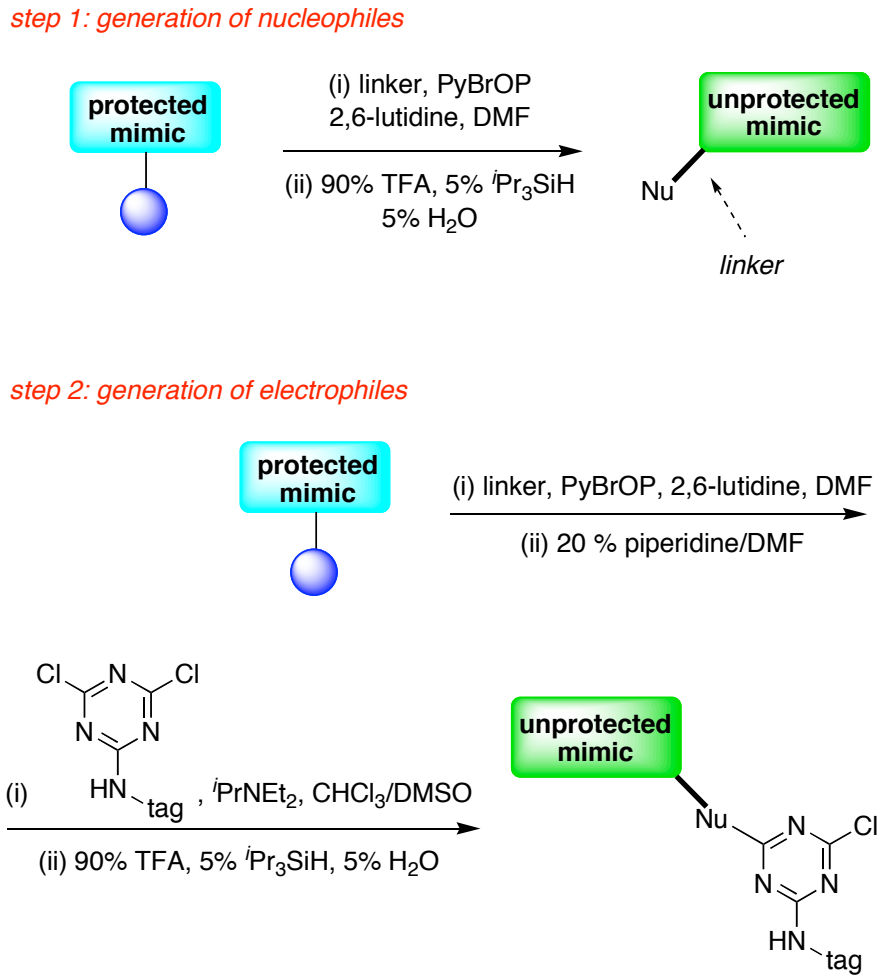
The pre-formed supported arylamine was divided into three portions and each was coupled with the short, medium or long linker individually with PyBrOP activation. Every portion with linker moiety attached was then split into two portions. One portion was subjected to FMOC deprotection and cleaved from the resin by treatment with a mixture of 90% TFA, 5% $HSi^{i}Pr_3$, and 5% H_2O. The crude peptide was triturated using anhydrous ethyl ether, re-dissolved and then purified via RP-Preparative HPLC to give monovalent **D3** derivative peptidomimetics. Another portion of linker-attached resin was FMOC deprotected first and coupled with dichlorotriazinylaminofluorescein (DTAF) or triethylene glycol (TEG) tag. After acid cleavage from the resin, the crude product was also purified by RP-preparative HPLC to yield the tagged peptidomimetics.

Assembly of the two monovalent molecules into one bivalent peptidomimetic was achieved by the method outlined in Scheme 2.2. This dimerization method has been established previously by our group,[43,69] however the work here is critically different insofar as the monovalent building blocks were essentially the same but the linker parts were varied as described above.

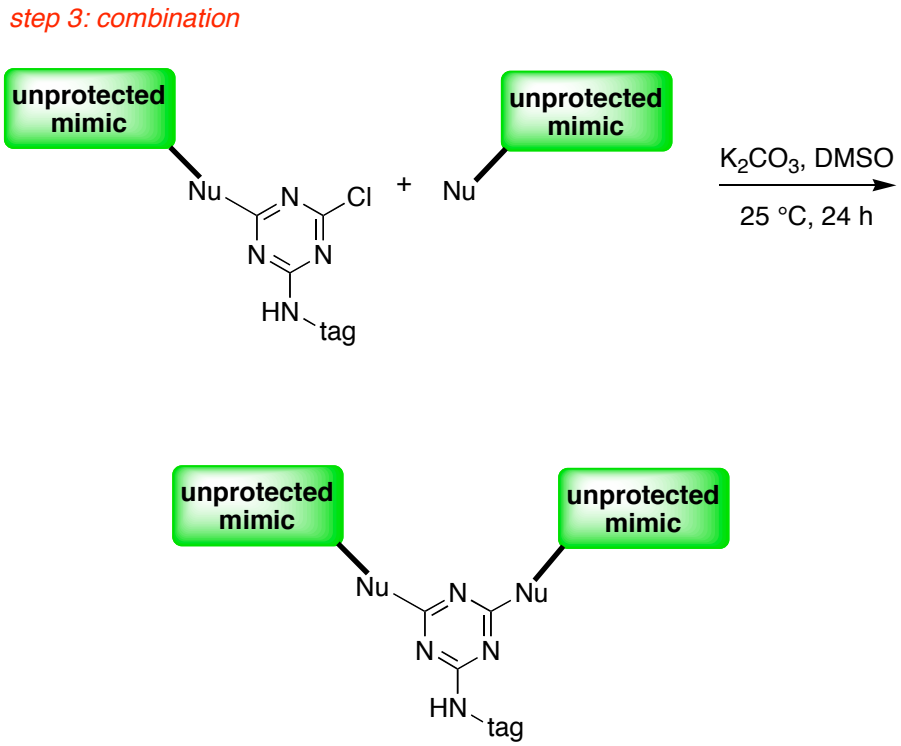
Firstly, stock solutions of the non-tagged fragment (nucleophile) and tagged fragment (electrophile) were prepared in DMSO. Then equal volumes of the two stock solutions of the non-tagged and the tagged peptidomimetic were mixed in the presence of solid K_2CO_3 as the base in 1 dr glass vials. The reaction vessels were capped and stirred at room temperature for 1-3 days. After HPLC analysis of the crude purity, the solution

in each vial was lyophilized to remove DMSO. The solid materials were re-dissolved in 1:1 mixture of H_2O/CH_3CN, and purified by RP-preparative HPLC to yield the final bivalent products.

Scheme 2.2. Overall scheme of preparation of the tagged bivalent peptidomimetics **1**.

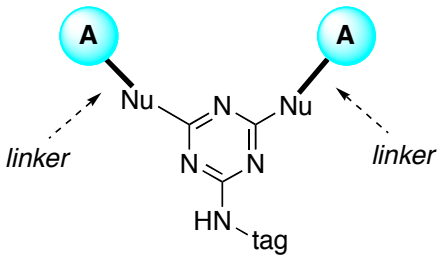


Scheme 2.2. continued.



From three linkers (**s**, **m** and **l**) and two tags a total of six bivalent compounds **1a-f** (**1-ss-F***, **1-sm-F***, **1-mm-F***, **1-sl-F***, **1-ml-F*** and **1-ll-F***) with fluorescein tag and three compounds with TEG tag **1g-i** (**1-ss-TEG**, **1-mm-TEG** and **1-ll-TEG**) were generated. Purities and compositions of the bivalent molecules were established by HPLC/MS are summarized in Table 2.1.

Table 2.1. Summary of the data of nine bivalent peptidimimetics **1a-i**.



labeled bivalent peptidomimetics **1a-i**

bivalent comp'd **1**	compound code	UV purity (%)[a]	retention time (min)[a]	amount (mg)[b]	$[M+H]^+$ found[c]
a	**1-ss-F***	100	10.567	3.1	1629.65
b	**1-sm-F***	100	11.050	2.0	1724.49
c	**1-mm-F***	100	11.267	3.9	1819.62
d	**1-sl-F***	98	12.683	4.0	1812.94
e	**1-ml-F***	100	12.917	4.8	1907.91
f	**1-ll-F***	100	14.133	3.0	1996.06
g	**1-ss-TEG**	100	9.267	5.1	1469.68
h	**1-mm-TEG**	100	10.133	2.5	1659.91
i	**1-ll-TEG**	100	12.967	2.5	1836.27

[a] Analytical HPLC analyses were carried out on 25 x 0.46 cm C-18 column using gradient conditions (10 – 90% B in 25 min) at 254 nm. The eluents used were: solvent A (H_2O with 0.1% TFA) and solvent B (MeCN with 0.1% TFA). The flow rate used was 1.0 mL/min.

[b] Calculated after preparative HPLC separation and lyophilization.

[c] Obatined from MALDI-MS.

2.4 Binding Assays

Binding assays were conducted by our collaborator, Dr. Uri Saragovi and his coworkers at McGill University in Canada. Direct binding assay and competitive binding assys were carried out for FITC-labeled peptidomimetics **1a-f**. Full details of analyses materials and methods are given in Appendix A.

2.4.1 Direct Binding Studies with Labeled Peptidomimetics

Direct FACScan-based binding assays were performed using FITC-labeled compounds and NIH-3T3 cells that were stably transfected to express the neurotrophin receptor TrkA, the neurotrophin receptor p75, or the control IGF-1R receptor to test for selectivity.

As shown in Figure 2.4, the use of longer length linkers based on the same triazine backbone caused a decrease in TrkA binding. Peptidomimetic **1-ss-F*** bound TrkA selectively and better than any other dimer. Progressively longer peptidomimetics **1-sm-F***, **1-mm-F***, **1-sl-F***, **1-ml-F***, and **1-ll-F*** exhibited significantly lower TrkA binding than **1-ss-F***. These studies were replicated using at least two independent syntheses of the compounds, with each synthesis being tested independently at least three times.

In additional specificity controls, NIH-3T3 cells that were transfected to express TrkC (NIH-TrkC) were also tested. In all cases they did not bind the compounds above the level for non-transfected cells (data not shown).

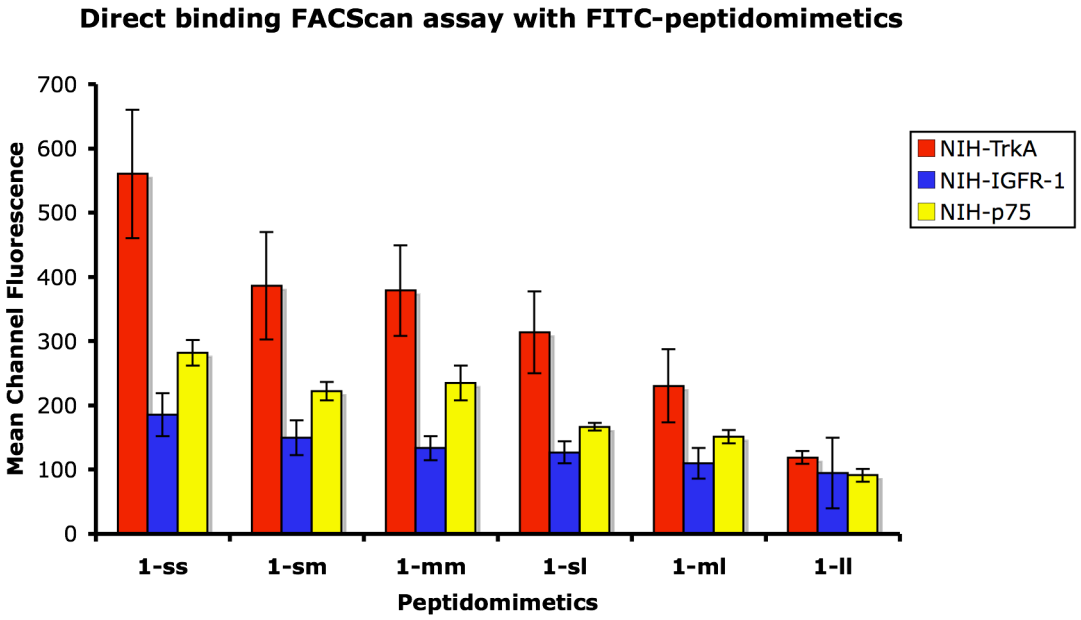


Figure 2.4. Direct binding FACScan assay with FITC-peptidomimetics. Cells expressing the indicated receptor were bound with test ligand (20 μM) at 4°C. After washing data was acquired and analyzed by FACScan/CellQuest with background subtracted. Mean channel fluorescence ± sem, n = 3-6 independent experiments.

2.4.2 NGF Binding Competitive Studies

Parental monovalent compound **D3** has been shown to bind TrkA in a way that does not compete with binding of NGF, and actually potentiates NGF activity. Wild type or TrkA-expressing NIH-3T3 cells was used to test the effect of **1-ss-F*** on the binding of ^{125}I[NGF] to TrkA (Figure 2.5). Unlabeled NGF was used as control competitor.

^{125}I[NGF] dose-dependent binding data subjected to Scatchard plot analyses showed the expected number of receptors on the cell surface and high affinity binding to TrkA (Figure 2.5a). A constant 500-fold excess of peptidomimetic **1-ss-F*** inhibited the binding of ^{125}I[NGF] by ~50% without changing the affinity of the residual ^{125}I[NGF] bound to TrkA. This result is suggestive of a competitive block. In a positive control, 100-fold excess cold NGF as competitor blocked nearly all the binding of ^{125}I[NGF]. The background cpms of these assays were determined by using the same concentration

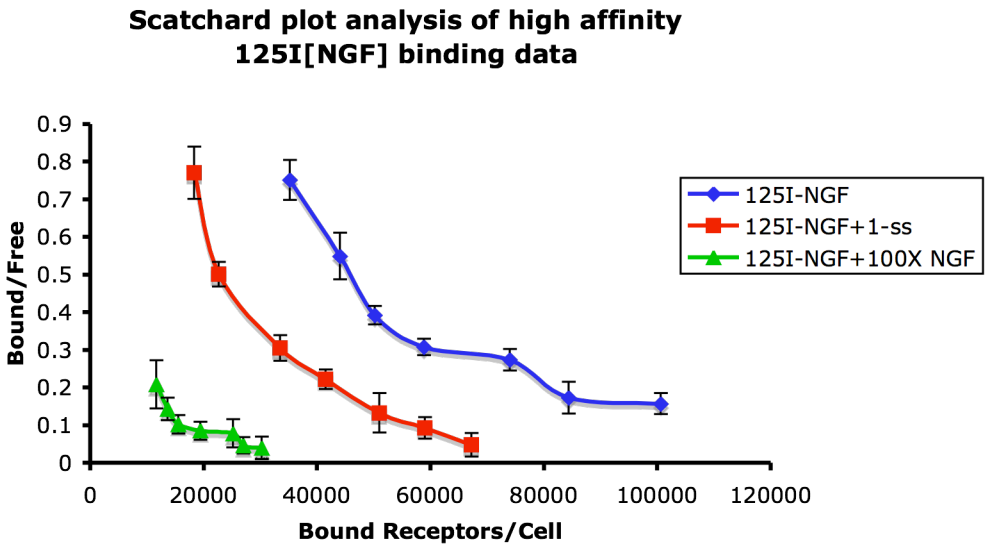
range of ^{125}I[NGF] binding to NIH-wild type cells (not expressing TrkA). In all assays, background ranged from 5-15% of the total binding, and in each case these cpms were subtracted.

Similar binding assays were carried out using a constant ^{125}I[NGF] concentration (1 nM, resulting in maximal 16,500 cpms bound) and a dose-range of inhibitors cold NGF or mimetic **1-ss-F***. The competitors were observed to reduce ^{125}I[NGF] binding in a dose-dependent manner. The IC_{50} of cold NGF competing ^{125}I-NGF is 2 nM. The IC_{50} of peptidomimetic **1-ss-F*** is 1,800 nM (~900-fold higher than NGF competing itself) (Figure 2.5b). From three different IC_{50} obtained for cold NGF *versus* peptidomimetic in independent assays, it was estimated that mimetic **1-ss-F*** has a $K_d \sim 135 \pm 30$ nM.

At last, anti-TrkA mAb 5C3 was used as a competitor because the parental mimetic **D3** competed with mAb 5C3 for binding. In this experiment MAb 5C3 competed for binding of FITC-labeled **1-ss-F*** in a dose dependent manner. The IC_{50} of mAb 5C3 upon mimetic **1-ss-F*** is ~60 nM (Figure 2.5c). Because the affinity of mAb 5C3 is ~4 nM,[70] we estimated that mimetic **1-ss-F*** has a $K_d \sim 240$ nM.

Overall, both estimates of affinity for **1-ss-F*** correlated well, and suggested higher affinity binding to TrkA than parental monovalent compound **D3**, which was anticipated from **1-ss-F*** bivalency.

a.



b.

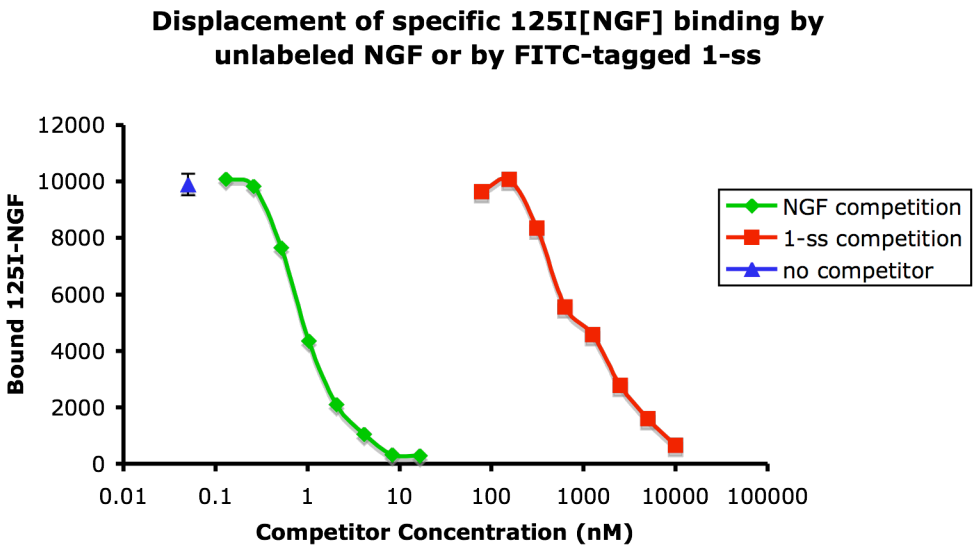


Figure 2.5. Binding competition studies using NIH-TrkA cells. **a.** scatchard plot analysis of high affinity ^{125}I[NGF] binding data; **b.** displacement of specific ^{125}I[NGF] binding by unlabeled NGF or by mimetic **1-ss-F***; **c.** mAb 5C3 competes for binding of FITC-tagged mimetic **1-ss-F***.

c.

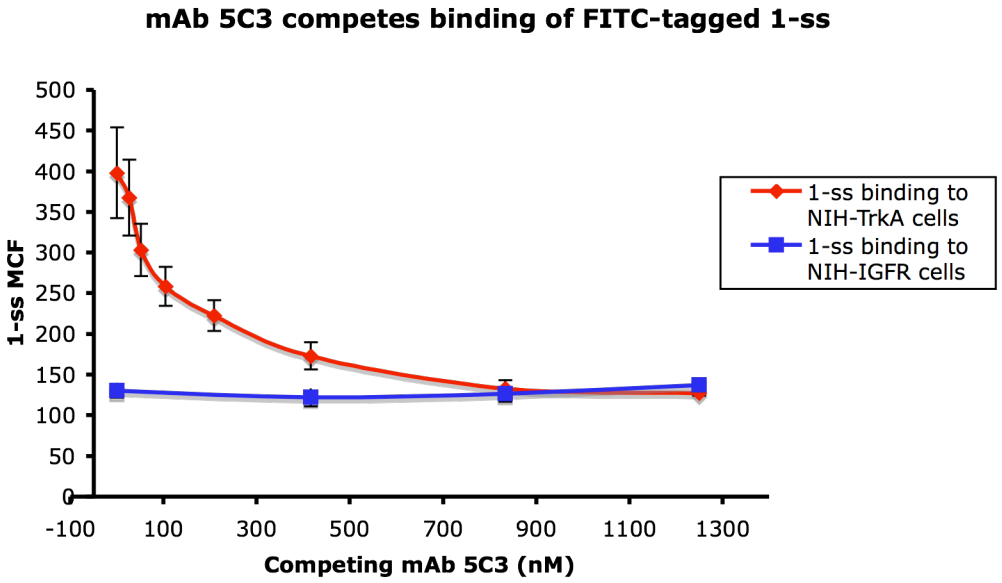


Figure 2.5. continued.

2.5 Receptor Dimerization Studies

The reason for increased affinity of **1-ss**, compared to parental monovalent A, is likely to be its bivalent binding. Moreover, blocking of NGF binding was only observed by **1-ss** but not by monovalent **A**, which may be due to the triazine linker bridge that fills the inter-receptor space where NGF fits. Both of these observations would be consistent with **1-ss** binding a TrkA-TrkA dimer. Thus, the chemical cross-linking experiments are performed to address this question.

NIH-TrkA cells were exposed to no ligand or to test ligands, followed by chemical cross-linking with disuccinimidyl suberate (DSS) to stabilize receptor dimers before resolution in denaturing SDS-PAGE and analyses by western blotting with highly specific anti-TrkA antibodies (Figure 2.6).

Untreated cells have a prominent band at p140 (cell surface TrkA monomer) and a less intense p110 band previously reported to be intracellular TrkA with a lower degree of glycosylation. There is also a less intense band at ~280 kDa detected in untreated

NIH-TrkA cells, which consistent with the existence of ligand independent pre-formed dimers in receptor over-expressing cells.[45] Exposure of NIH-TrkA cells to NGF (10 nM), to **1-ss-F***, or to **1-ss-TEG** increased formation of TrkA dimers at 280 kDa. Without chemical cross-linking, there is no detection of bands at ~280 kDa because the receptor dimers are non-covalent (data not shown).

These data indicate that **1-ss** and NGF each is able to induce or to stabilize TrkA dimers at the cell surface, in such a way that covalent chemical cross-linking at the cell surface is more efficient. These data support the notion that one bivalent **1-ss** molecule may bind two TrkA receptors.

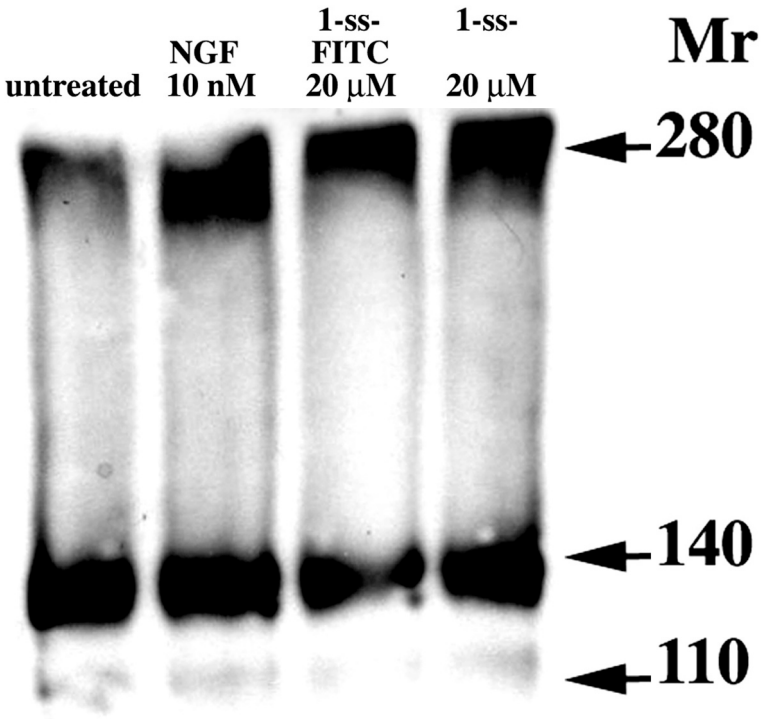


Figure 2.6. Induction of putative TrkA dimers. After exposure to the indicated ligand (30 min at 4°C), NIH-TrkA cells were chemically cross-linked with DSS. After washing, they were detergent solubilized and samples (20 mg protein/lane) were studied by western blotting with a highly specific anti-TrkA mAb 5C3.

2.6 Cellular Assays

Biological assays were also conducted by Dr. Uri Saragovi and his coworkers at McGill University in Canada. These assays were carried out especially for TEG-labeled peptidomimetics **1g-i**. Full details of analyses materials and methods are given in Appendix A.

2.6.1 Cell Survival Assays

Using NIH-3T3 cells expressing either TrkA, TrkC or IGF-1R in quantitative MTT assays, the effect of peptidomimetics on bioactivity was evaluated in functional studies of receptor-mediated cell survival. The cells undergo apoptosis when they are placed in serum-free media (SFM). In these conditions, cells can be protected by their appropriate growth factors (NIH-TrkA is protected by NGF, NIH-TrkC is protected by NT-3, NIH-IGF-1R is protected by IGF-1). Growth factor protection of cells from apoptosis in SFM is dose-dependent; suboptimal doses of growth factor quantitatively and consistently give reduced survival (25-40% of maximal).

In initial binding assays the FITC-containing peptidomimetic **1-ss-F*** was used and these assays proved that **1-ss-F*** was an antagonist of NGF (data not shown). However, it is not desirable to use a FITC-labeled compound in bioassays. To verify the irrelevancy between the FITC moiety and the antagonistic function, we tested analogs **1-ss-TEG**, **1-mm-TEG**, and **1-ll-TEG** that contain a TEG instead of the FITC moiety in additional survival assays.

Mimetic **1-ss-TEG** showed antagonistic function of NGF bioactivity, measured as NGF-mediated protection of NIH-TrkA cells stressed to undergo apoptosis (Figure 2.7a). Antagonism was selective because **1-ss-TEG** did not antagonize the protective functions of NT-3 upon NIH-TrkC cells (Figure 2.7b), or the function of IGF-1 upon NIH-IGF-1R (data not shown). In further control assays, mimetics **1-mm-TEG** and **1-ll-TEG** had no significant effect on the survival-promoting effect of any growth factor, NGF (Figure 2.7a), NT3 (Figure 2.7b), or IGF-1 (data not shown). Although inactive, these bivalent mimetics are highly related to the active mimetic **1-ss-TEG** and are therefore controls of great importance.

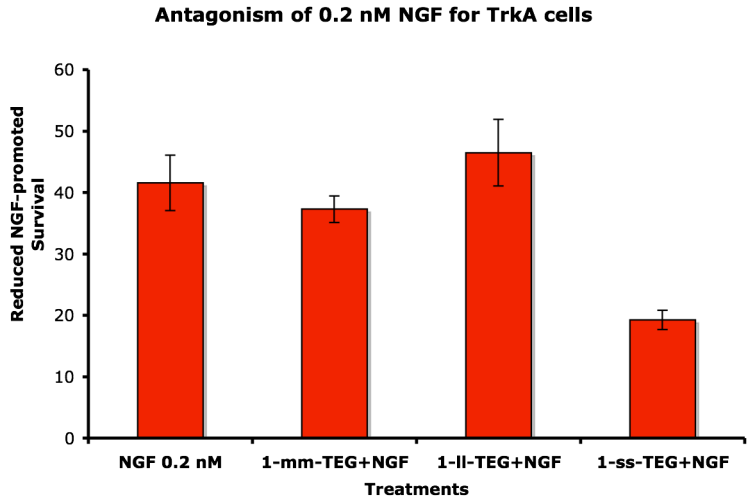
Antagonism of NGF action by **1-ss-TEG** was found to be dose-dependent. The survival promoted by 0.2 nM NGF (~40% of maximal survival) was reduced by **1-ss-TEG**. The mimetic at 5 μM, 20 μM, and 50 μM respectively reduced the survival promoted by NGF from ~40% to ~25% (non-significant), and to a statistically significant ~15%, and ~0% survival (Figure 2.7d).

Interestingly, mimetic **1-ss-TEG** also significantly accelerated the death of NIH-TrkA cells in SFM (in the absence of NGF) (Figure 2.7c). Over-expressed TrkA in these cells was known to have spontaneous ligand-independent trophic functions. In the experiment high baseline ligand-independent activity was observed in cells over-expressing TrkA and other RTKs.

Antagonism of baseline TrkA activity by **1-ss-TEG** was also TrkA selective, for the fact that inhibition of ligand-independent activity was not observed in NIH-TrkC cells (Figure 2.7c) or NIH-IGF-1R cells (data not shown). Moreover, accelerated death of NIH-TrkA cells in SFM by **1-ss-TEG** was dose dependent and significant at concentrations of 20 μM and 50 μM (Figure 2.7e).

In control assays, mimetics **1-mm-TEG** and **1-ll-TEG** showed no significant effect on ligand-dependent activation (Figure 2.7a and b), or in baseline ligand-independent activity (Figure 2.7c) of any receptor. And there was no detectable toxicity in MTT assays when cells were cultured in serum containing media + 50 μM of the mimetics (data not shown).

a.



b.

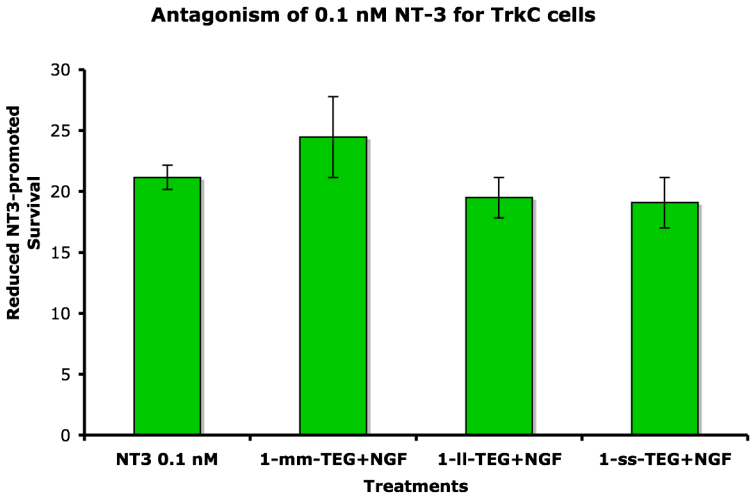
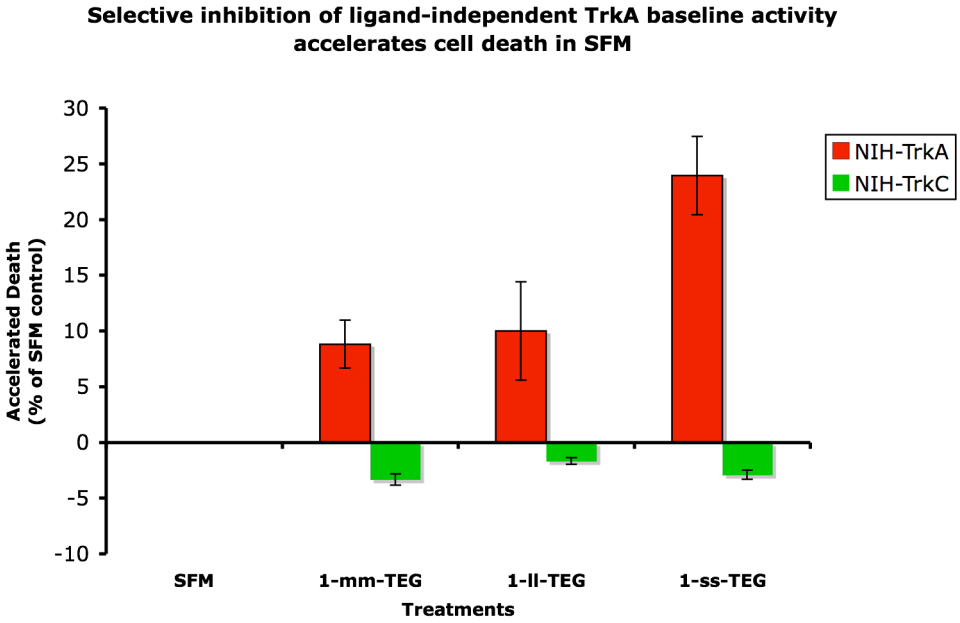


Figure 2.7. Mimetic **1-ss-TEG** antagonizes NGF and TrkA in cell survival assays. Mimetics were tested at 20 μM for antagonism of **a.** 0.2 nM NGF for TrkA cells; **b.** 0.1 nM NT-3 for TrkC cells; or **c.** accelerated cell death in SFM. Antagonism of **d.** NGF; or **e.** TrkA baseline activity were dose-dependent.

c.



d.

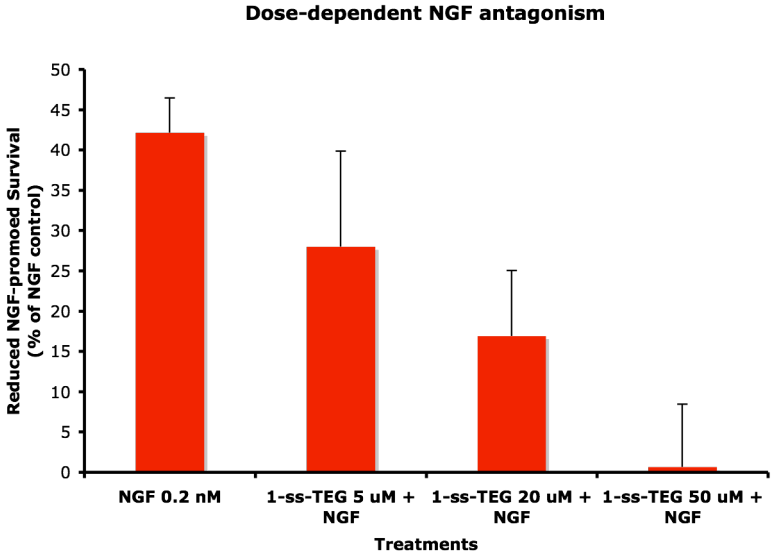


Figure 2.7. continued.

e.

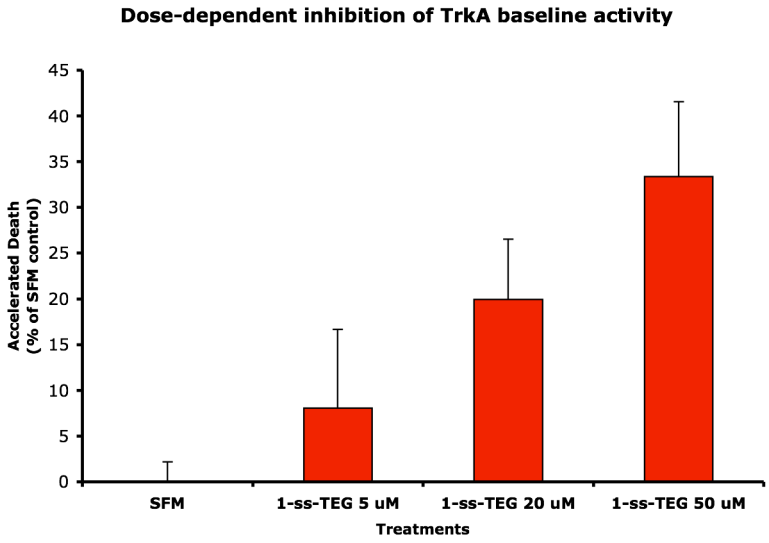


Figure 2.7. continued.

2.6.2 Antagonism of Biochemical Signals by 1-ss

Over-expressed TrkA in these cells has spontaneous ligand-independent trophic functions; and high baseline ligand-independent activity has been reported in cells over-expressing TrkA and also for other RTKs. Indeed, the presence of TrkA-dimers was observed in these cells in the absence of ligand (e.g. see Figure 2.6).

For this reason, we followed the observation that mimetic **1-ss-TEG** significantly accelerated the death of NIH-TrkA cells in SFM in the absence of NGF. We hypothesized that this may be due to antagonism of baseline receptor activity, which is partly protective of cell death. We carried out biochemical analyses of receptor-tyrosine phosphorylation (pTyr) (Figure 2.8) by western blotting with anti-phosphotyrosine antibodies untreated or ligand treated cells.

Untreated cells have some TrkA-pTyr, which decreases significantly after treatment with 1-ss-TEG (10 μM). This was seen after cells were treated with ligand for 2 hours (Figure 2.8 lane 1 versus 2) or 20 minutes (Figure 2.8 lane 5 versus 6). As expected, control treatment with NGF (10 nM) for 2 hours (Figure 2.8 lane 3) or 20 minutes (Figure 2.8 lane 7) results in significant increases in TrkA-pTyr. Interestingly, **1-**

ss was also able to reduce the level of TrkA-pTyr even in the presence of NGF (Figure 2.8 lane 4). To show equal receptor loading, the membranes were stripped and re-blotted with anti-TrkA mAb 5C3.

All these data are consistent with the results of the bioassays, and indicate that **1-ss** has the ability to inhibit the baseline kinase activity of TrkA. It is curious that **1-ss** can do so while at the same time it can also increase the number of TrkA-TrkA dimers (e.g. see Figure 2.6), because dimeric receptors are thought to be in the active state.

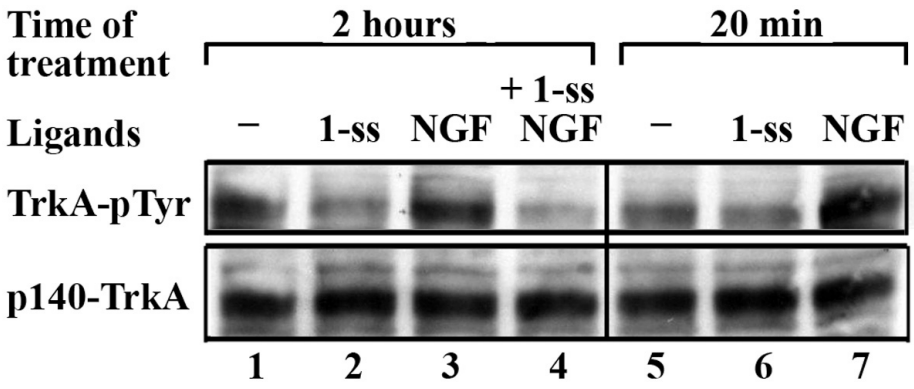


Figure 2.8. Mimetic **1-ss-TEG** antagonizes baseline (ligand-independent) and ligand-dependent TrkA activity. NIH-TrkA cells were cultured in SFM alone, and were then supplemented with the indicated ligand for 2 hours or for 20 min at 37°C in a cell incubator. After detergent solubilization the phosphotyrosine levels of the samples were studied by western blotting with anti-pTyr mAb 4G10. Total TrkA loading was verified on the same membranes with a highly specific anti-TrkA mAb 5C3.

2.7 Docking/Modeling

It is interesting to notice that the previous result of binding assays and biological assays of parental monovalent compound **D3** showed that it bound to TrkA in a non-competitive manner with binding of NGF and it potentiated NGF activity.[50] However bivalent peptidomimetic **1-ss** seems not to follow the same binding mode with NGF from the data shown in the previous section. To better understand and explain the basis for NGF competition and receptor antagosim of **1-ss,** molecular modeling was carried out to

explore the interaction mode. This modeling was conducted by Dr. Uri Saragovi and his coworkers at McGill University in Canada. More details of docking/modeling methods are given out in Appendix A.

Firstly, the 3D structures of **1-ss** were built using SYBYL optimized by using Tripos force field[71] and partial charges of **1-ss** were calculated by the AM1-BCC method.[72] The 200 lowest-energy conformers of **1-ss** were selected.

A model of the target TrkA-D5 was also prepared. A rectangular box was constructed following the NGF-TrkA-D5 crystal complex (Figure 2.9a and 2.9b); [73,74] this box defined the volume to be explored in docking **1-ss** to TrkA-D5.

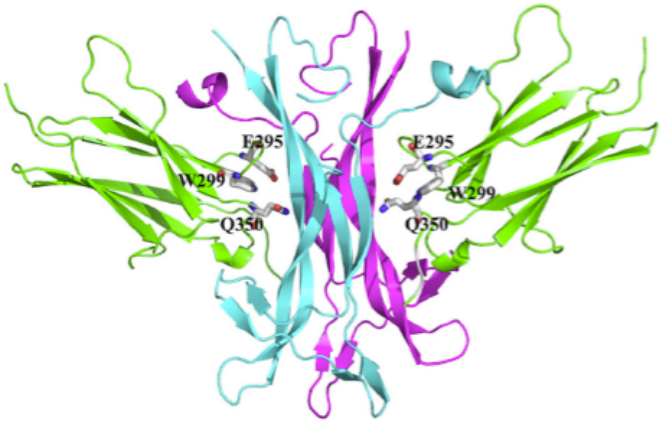
These two molecules were then docked in an unbiased manner to explore possible binding sites. The possible binding modes for the 200 lowest-energy conformers of **1-ss** were posed on the TrkA-D5 rectangular box using a docking algorithm, WILMA; and resulted in 52 clusters of bound conformers. The top-scoring poses of **1-ss** had similar docking scores (–6 to –9 kcal/mol). One of the top-scoring poses is shown in Figure 2.9c and 2.9d, in which **1-ss** binds simultaneously to two monomers of TrkA-D5. Each pharmacophore of the **1-ss** homodimer binds to one monomer of the TrkA-D5 dimer, occupying mainly the membrane proximal part of TrkA-D5, which is solvent exposed. This is consistent with our hypothesis of non-covalent cross-linking of the receptor pair by **1-ss**.

A close-up of the binding mode of this pose is also shown in Figure 2.9d. The lysine side-chain of one monomer of **1-ss** forms a salt bridge (2.6 Å) with the carboxylate group of E295 of TrkA-D5. The glutamic acid side-chain of the same monomer also forms weak H-bonding interactions with the backbone amide of the Q350 (3.8 Å) and indole nitrogen of W299 (2.8 Å) of TrkA-D5. The linker of **1-ss** is mostly solvent exposed, but the amides connected to the triazine can form two weak H-bonds with N349 (3.0 Å) and Q350 (3.0 Å) of the contralateral TrkA-D5 monomer. Therefore, the linker may contribute, although minimally, towards receptor binding.

Besides, the binding affinity of this pose was estimated using a solvated interaction energy (SIE) approach[75] and a value of -8.87 kcal/mol (0.29 μM) was obtained. This value was compared with the experimental binding affinity (~0.2 μM) of

1-ss and found to be close, which further proved the consistency of the proposed binding mode with the experimental binding affinity and observed NGF binding inhibition.

a.



b.

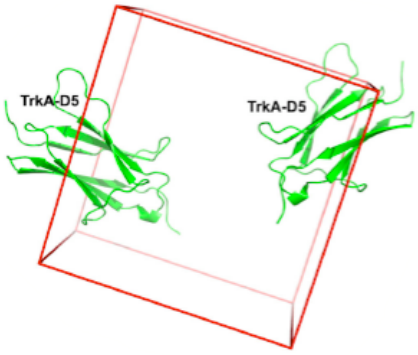
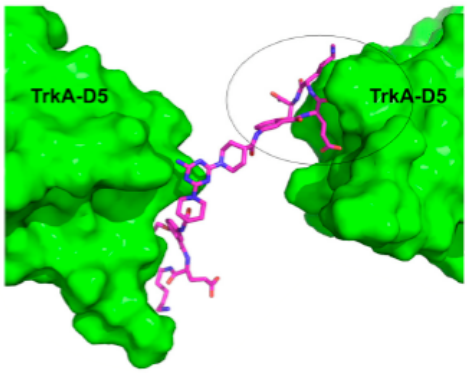


Figure 2.9. Molecular modeling illustration of NGF:TrkA and **1-ss**:TrkA interactions. **a.** ribbon diagram representation of TrkA dimer (green) bound to NGF dimer (cyan and light blue). The capped stick amino acids of the receptor indicate where monovalent **D3** binds, near the NGF hot spot; **b.** the red box representing the putative binding pocket of compound **1-ss,** used for docking; **c.** the surface of TrkA-D5 (green) with **1-ss** shown in capped stick representation; **d.** a close-up of the circled area in **c** as representation of TrkA-D5 (green) complexed with **1-ss**. The residues of TrkA-D5 forming hydrogen bonds with **1-ss** (original atom color) were represented by capped stick models. The hydrogen bonds between **1-ss** and TrkA-D5 are shown by dashed lines.

c.



d.

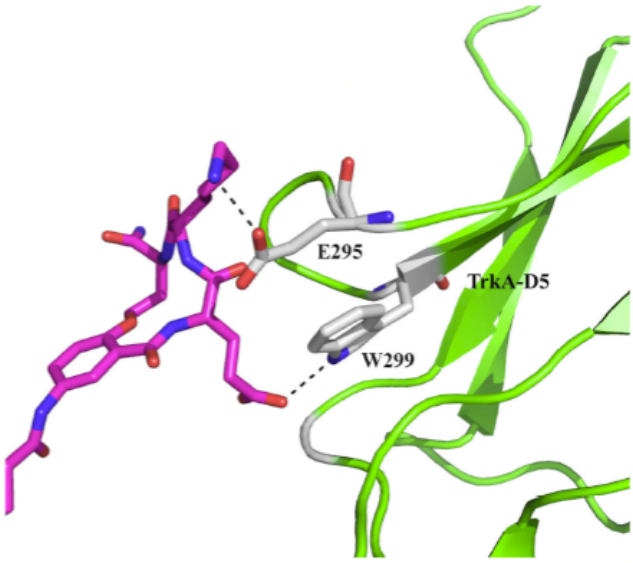


Figure 2.9. continued.

2.8 Discussion and Summary

An early lead, cyclic monovalent peptidomimetic **D3** was found to bind NFG receptor, TrkA selectively and have intrinstic agonistic activity in the absence of NGF based on previous studies. To improve the binding affinity, a library of nine new bivalent peptidomimetics that features two close monovalent moieties with **D3** and linker combinations of various lengths (**ss**, **sm**, **mm, sl, ml, ll**) was prepared via Fmoc-SPPS,

S_NAr macrocyclization reactions and a highly efficient dimerization method. A fluorescein tag or TEG tag was attached to each dimer to facilitate the binding assay or biological assay.

Direct binding assay and NGF competitive assay showed that **1-ss** has higher affinity than other bivalent compounds with TrkA and a ~50-fold better affinity than the parental monovalent peptidomimetic **D3** (~200 nM *versus* ~10 μM), and retains high selectivity towards TrkA. While both molecules can block the binding of another ligand, anti-TrkA mAb 5C3 (the latter is more efficient), **1-ss** can also block NGF binding. In the subsequent biological assays, **1-ss** was found to antagonize ligand-dependent activation of TrkA as a pure receptor antagonist while the parental monovalent peptidomimetic **D3** is an agonist towards NGF.[50]

It is intriguing that the parental monovalent molecule **D3** potentiates or facilitates NGF function (logically does not block NGF binding) and dimeric **1-ss** with higher avidity and affinity to a different receptor hot spot blocks the binding and bioactivity of NGF. The observation that **1-ss** increases TrkA-TrkA dimers suggests that this bivalent mimetic may bind to two receptors; this is also consistent with the docking result that **1-ss** binds two TrkA receptors. This model suggests that **1-ss** can block NGF binding through the extended linker that bridges the homodimer and "fills" the inter-receptor space where NGF normally fits, which is probably why **1-ss** is a competitive antagonist of NGF without binding to the same receptor hot spot. Logically, blocking NGF binding also results in antagonism of NGF's trophic activity as well. Mimetic **1-ss** reduces ligand-independent receptor activation in trophic survival assays, and reduces the baseline tyrosine phosphorylation of TrkA.

The functional difference between **1-ss** and **D3** can be further interpreted by their different ability to allow or disallow putative receptor conformational changes that ought to take place before or after binding of an agonist to TrkA. However, this is speculation, and further research is required to address whether dynamic conformational changes of the TrkA receptor may be important for bioactivity.

To our knowledge, this is the first report of an agonist being converted to an antagonist through simple dimerization. The experience can provide strategies to

discover inhibitors or activators of NGF receptors of new classes, and could possibly be expanded to other receptors.

Overall, in this work we address whether and how dimerization of an agonistic monovalent ligand could generate an antagonist, whether simple binding to an agonistic hot spot of the receptor TrkA was sufficient for activation, or whether receptor conformational changes would be required to afford agonism.

From the assay result and computer modeling, we propose that if receptor conformational changes must be induced when it binds an agonistic ligand, a relatively rigid bivalent ligand could "freeze" the receptor dimer by non-covalently cross-linking the receptor pair and prevent conformational changes. Unfortunately, conformational changes in TrkA upon the binding of its protein ligand NGF have not yet been demonstrated through structural analyses, and this is still an issue in the current debate.[54]

CHAPTER III

A COMBINATORIAL METHOD TO SOLUTION-PHASE SYNTHESIS OF BIOTIN-LABELED BIVALENT BETA-TURN MIMICS

3.1 Introduction

2,5-Diketopiperazines (DKPs), an important class of heterocyclic compounds, are the smallest cyclic peptides and have been known for over a century.[76-78] They have attracted a lot of attention in organic and medicinal chemistry for their synthesis and biological activities.[79-83] This unique heterocyclic system has been found in several natural products and many interesting biologically active synthetic molecules.[84-88] Their wide spectrum of biological properties have gained significance in modern drug discovery.[89]

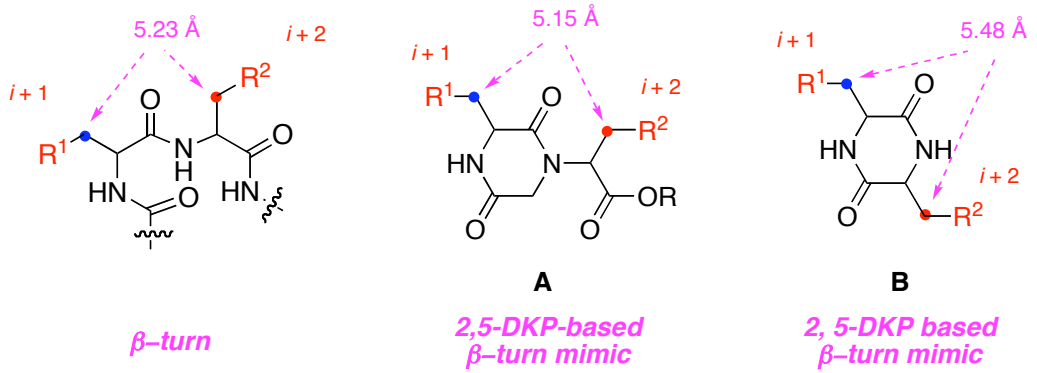
Biologically, DKPs are related to some important bioactivities, such as inhibition of various enzymes, as well as plasminogen activator inhibitor-1,[90-93] opioid receptor agonism and antagonism,[94] and alternation of cardiovascular and blood-clotting functions.[82] They also exhibit activities as antitumor, antiviral, antifungal, antibacterial, antihyperglycaemic agents, etc.[79,80] These, including their resistance to proteolysis, make them very attractive for medicinal chemistry.

Chemically, DKPs show well-defined and controlled substituent group stereochemistry in up to 4 combinations, which can provide structural diversity with particular orientation; this also makes them privileged structures for combinatorial chemistry in lead discovery.[29,95] In addition, their conformational rigidity, and availability of hydrogen bonding donor and acceptor groups, is greatly favored in various interactions with biological targets. Thus DKPs are frequently involved in mimicking of peptidic pharmacophoric groups,[96,97] and their favorable characteristics are considered ideal for the rational development of new drugs.[82,98]

Synthetically, DKPs are a common scaffold that can be accessed easily from conventional procedures in good yields.[79,80] Many DKPs can be synthesized efficiently from α-amino acid esters in solution, via intramolecular cyclization of corresponding dipeptides;[28] this method does not require many reaction steps. Dipeptides that are fixed on the solid support via a bond that is susceptible to nucleophilic attack (*e.g.* phenacyl ester bond),[98] or even better, a safety-catch linker bond that is more stable and activated to allow for cleavage and cyclization,[99] are commonly used building blocks in solid-phase DKP synthesis. Another way to obtain DKPs is via Ugi reaction, a multi-component reaction also known as the Ugi 4-component coupling (4-CC).[100-103] This approach is ideal for generation of highly substituted DKP libraries by parallel solution-phase synthesis.[104] Besides these conventional methodologies, microwave heating as an innovating strategy was used to synthesize symmetrical dimeric DKPs without epimerization at α-position of the amino acids in higher yields and shorter reaction times.[105]

Most of the known DKPs generated from above procedures have a general structure as **B** (Figure 3.1a), the two substituents are attached to the α-carbons of the two amino acids; these side-chains usually carry the functionalities.[28,106] However, during our attempts to overlay the optimized target structure onto a standard type I β-turn motif, we found that it did not fit very well with the backbone (Figure 3.1b). This led to structural modification and development of a better design, which is shown as structure **A** (Figure 3.1a). In DKP **A**, the two side-chains from amino acids were retained, however at more favorable positions to offer a better overlay with standard β-turn in molecular modeling (Figure 3.1b). Of course, this changes the synthetic route correspondingly.

a.



b.

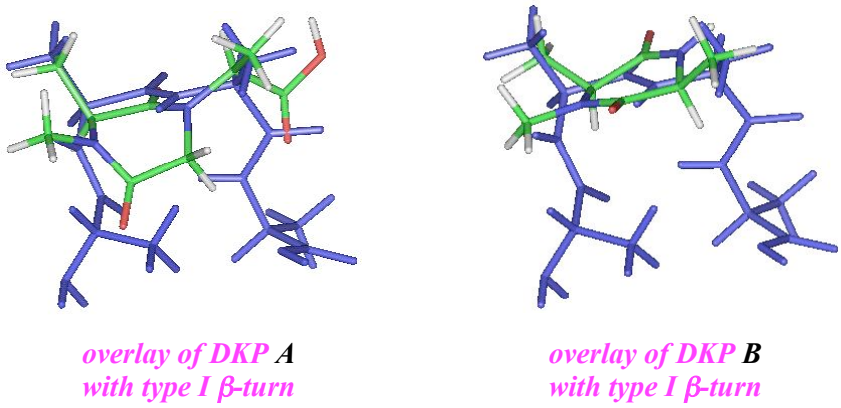


Figure 3.1. a. The key distance of C^{b}-separations of the i + 1 to i + 2 residues of a type I β-turn and of the monovalent turn mimics DKP **A** and **B** featured here; **b.** an overlay comparison of the two types of 2,5-DKP mimics onto a type I β-turn.

Dimerization of bioactive molecules is a strategy that is used to enhance the binding affinity by introducing more pharmacophores around the recognizing sites. Weakly binding molecules could also be jointed together to form bivalent molecules with higher binding affinity. Actually a simple DKP, *i.e.* cyclo(Gly-Gly), itself has been used to interlink two other molecules.[107] In this work, we followed a very efficient combinatorial methodology to generate a library of bivalent DKP-based peptidomimetics in solution.[43,69] Two different length linkers (short and long) were used to increase the number of permutations, and more importantly, explore the optimal linker dimensions for

dimeric ligands docking with dimeric receptors. A biotin derivatized 'tag' was attached to each bivalent molecule to facilitate further binding and biological assays.

3.2 Preparation of Monovalent Diketopiperazine Peptidomimetics

In attempts to prepare the monovalent diketopiperazine building blocks, two routes were developed. From four different DKPs and two linkers (short and long), a total of eight monomeric compounds were constructed. The synthetic route (Scheme 3.1-a) was chosen based on the amino acid sequence that was incorporated.

For cyclo(Gly-Lys) and cyclo(Thr-Gly) (Scheme 3.1a), the synthesis began with Boc-protected *N*-methyl amino acid and *N*-terminal free amino acid *tert*-butyl ester. They were coupled under typical peptide coupling conditions. The Boc group was then selectively removed[108] in the presence of *tert*-butyl ester using TFA. S_N2 reaction between bromoacetyl bromide and the deprotected secondary amine intermediate[109] generated the linear bromide precursor. Using a strong base, the NH from the backbone amide was deprotonated and nucleophilic attack at the carbon next to bromide[109] produced the cyclization product **2a-b** as a 2,5-diketopiperazine ring. After TFA removal of the *tert*-butyl ester[108] we can obtain the corresponding carboxylic acid **3a-b**, which is the handle for linker attachment.

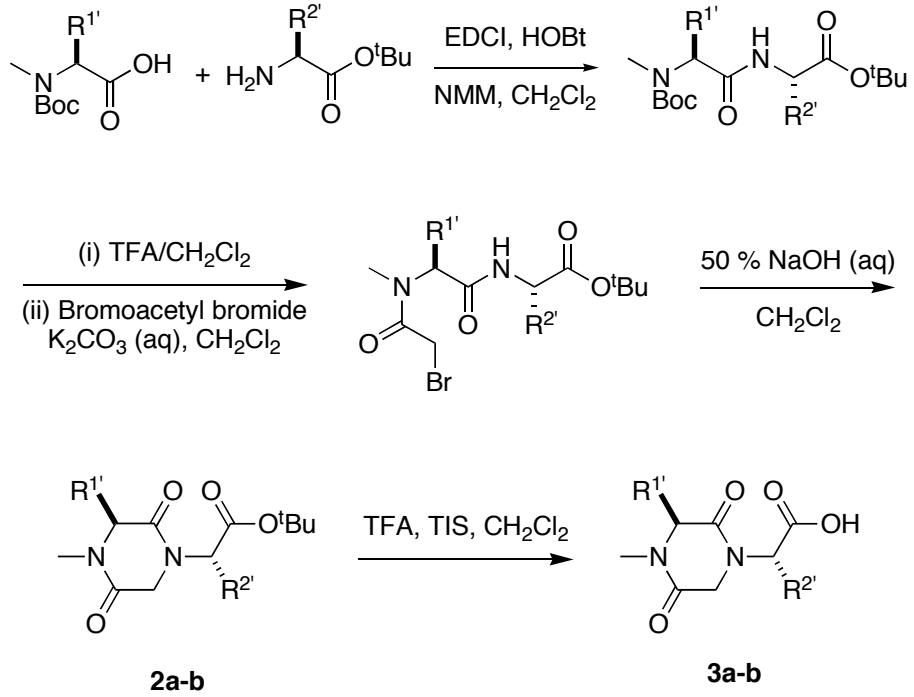
For cyclo(Glu-Lys) and cyclo(Ile-Lys) (Scheme 3.1b), the synthesis was slightly different. It started from Cbz-protected *N*-methyl amino acid and *N*-terminal free amino acid methyl ester coupling as above. Accordingly, the dipeptide was deprotected by hydrogenolysis and the free secondary amine intermediate reacted with bromoacetyl bromide. The cyclization condition remained the same, however a gentle hydrolysis condition to remove the methyl ester in **2c-d** was needed without causing racemization. After several tests, K. C. Nicolaou's condition[110] (trimethyltin hydroxide) was found to be ideal for this case to give the enantiomerically pure diketopiperazine carboxylic acids **3c-d** in good yields.

With four DKP carboxylic acids **3a-d** in hand, the next, as well as the last step for cyclo(Glu-Lys) and cyclo(Ile-Lys) monomers, was the linker coupling (Scheme 3.1c). Each carboxylic acid was efficiently coupled with a short or long linker to generate eight

monovalent compounds **4a-d** and **5a-d** in total. For cyclo(Gly-Lys) **4a** and **5a,** and cyclo(Thr-Gly) **4b** and **5b** monomers, side-chain protections (Cbz, Bzl) were finally removed via convenient hydrogenolysis reactions to offer **6b** and **7b**. For cyclo(Gly-Lys), a Boc protection[111] was installed back on the side-chain to give **6a** and **7a** for easier purification and handling. The overall syntheses were successful with decent yields and all eight desired products were purified via flash column chromatography before being subjected to dimerization reactions. Scheme 3.1 uses $R^{1'}$ and $R^{2'}$ to denote protected side-chains (and R^1 and R^2 to indicate deprotected ones).

Scheme 3.1. Two methods for preparing the monovalent mimics **4** and **5**. Route **a.** featuring cyclo(Gly-Lys) and cyclo(Thr-Gly) monomers; Route **b.** featuring cyclo(Glu-Lys) and cyclo(Ile-Lys) monomers; **c.** linker attachment (for **4a-b** and **5a-b,** there is an additional hydrogenolysis step afterwards using H_2 and Pd-C in MeOH to remove side-chain protections to give **6b** and **7b**; and a Boc protection in CH_2Cl_2 to give **6a** and **7a**).

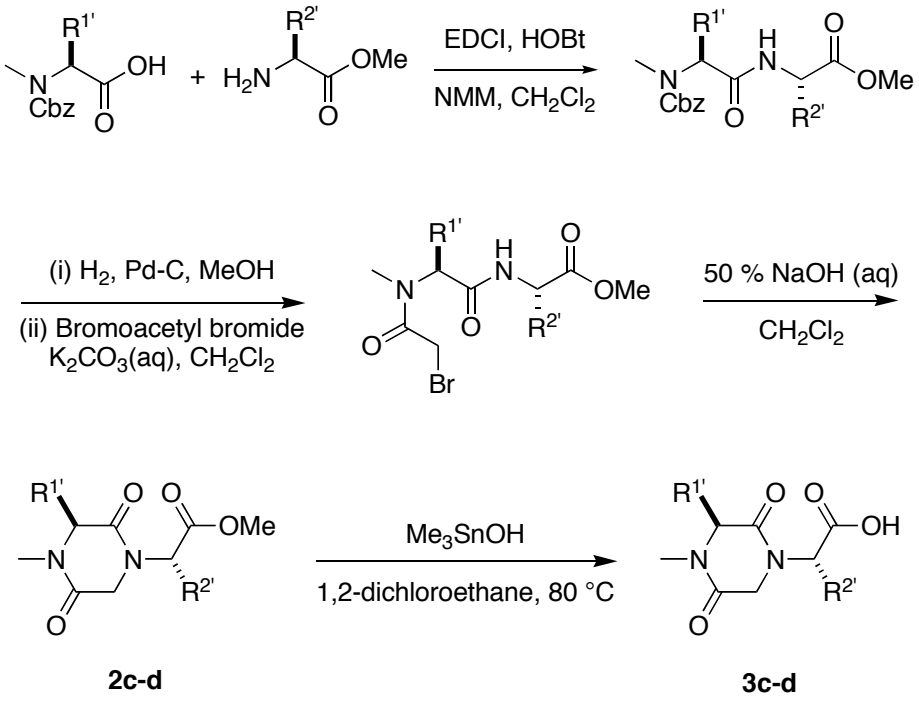
a.



2a-b **3a-b**

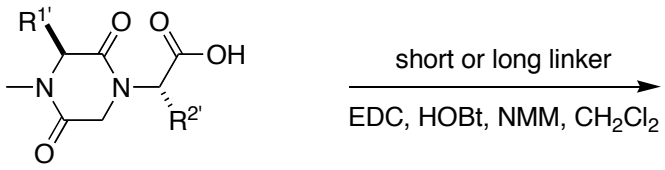
Scheme 3.1. continued.

b.

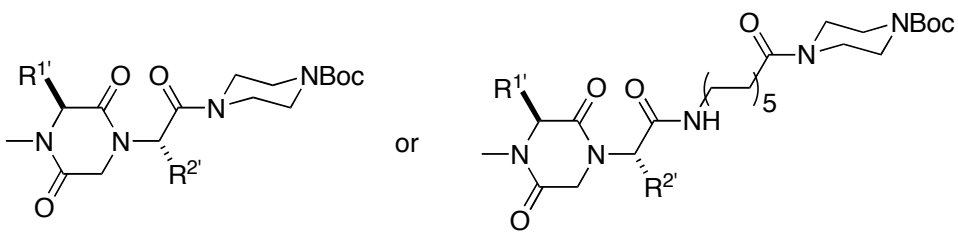


2c-d **3c-d**

c.



3a-d



4a-d

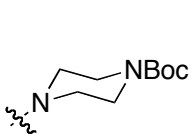
short linker attached

5a-d

long linker attached

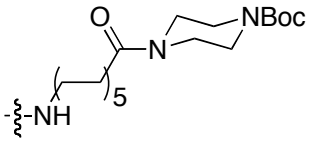
The data of eight monomeric diketopiperazine compounds prepared via methods **a** and **b** described above were summarized in Table 3.1.

Table 3.1. Monovalent peptidomimetics **4c-d**, **5c-d**, **6a-b** and **7a-b** prepared via method **a** and **b**.



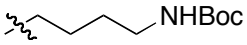
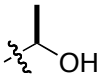
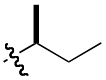
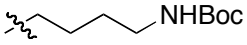
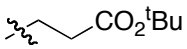
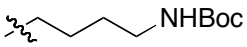
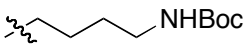
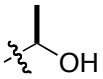
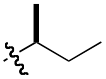
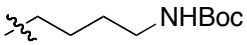
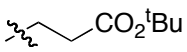
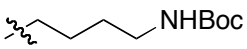
I

short linker



II

long linker

compound	$R^{1'}$	$R^{2'}$	linker	yield (%)[a]
6a	H		I	48
6b		H	I	61
4c			I	57
4d			I	48
7a	H		II	49
7b		H	II	50
5c			II	55
5d			II	40

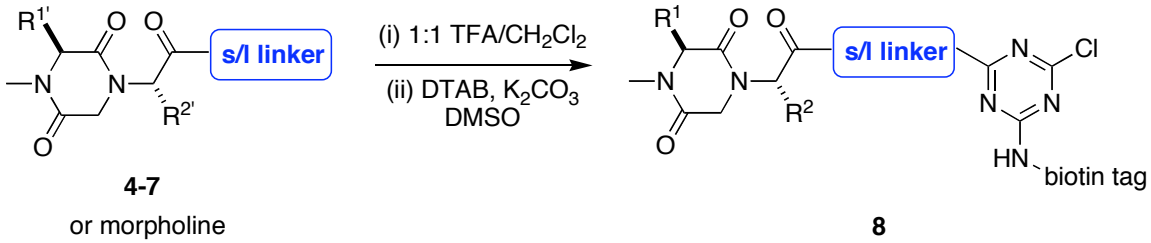
[a] From intermediate **3a-d** after flash chromatography.

3.3 Application of the Solution-phase Procedure to Make a Library of Biotin-labeled Bivalent Mimics

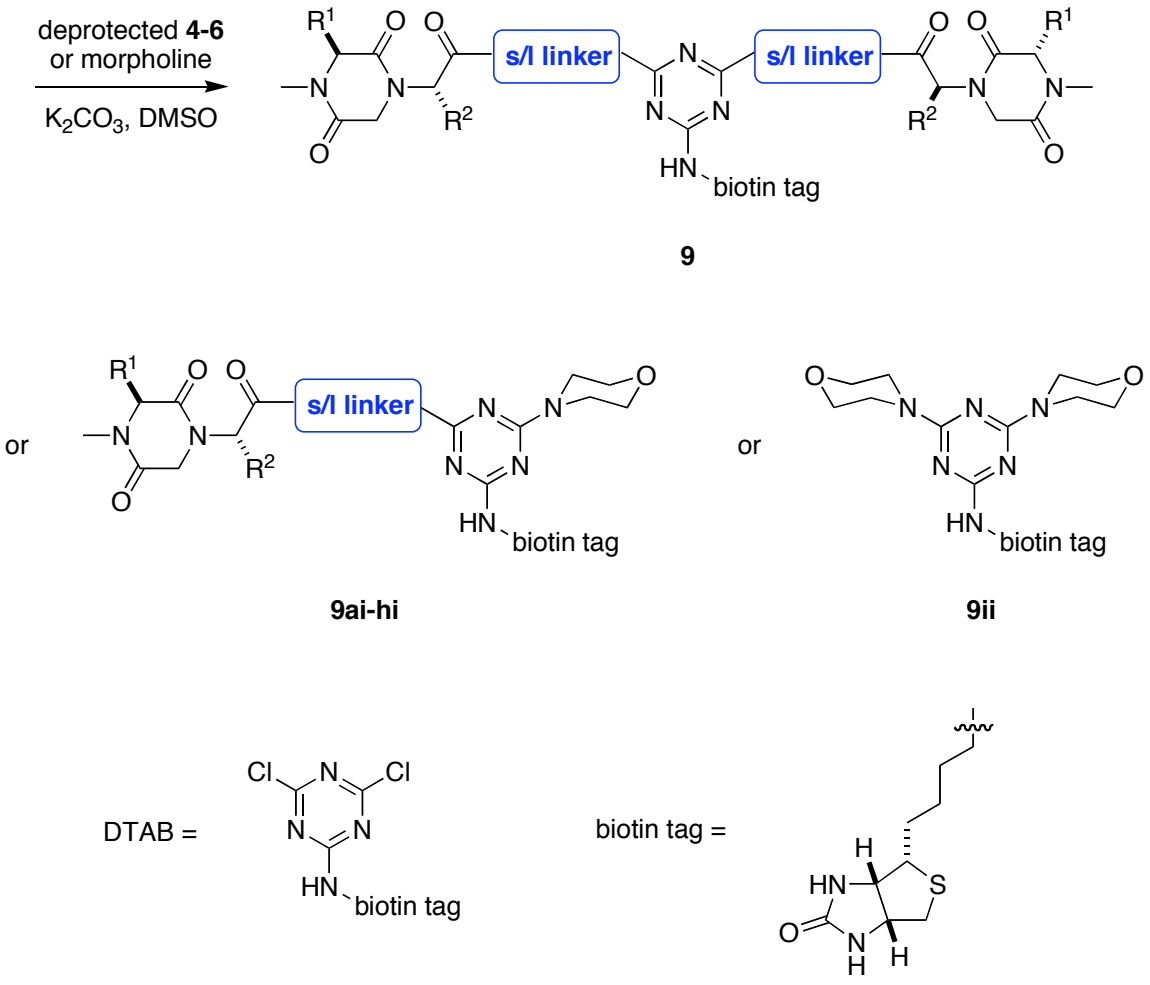
The protected monomeric DKPs **4c-d**, **5c-d**, **6a-b** and **7a-b** were subjected to global deprotection to give unmasked mimics then reacted with a dichlorotriazine derivative of biotin (*i.e.* DTAB) to give the electrophiles **8**. These were then reacted with second aliquots of the mimics **4c-d**, **5c-d**, **6a-b** and **7a-b**, in all permutations, to give the library of bivalent mimics **9**. Morpholine was also used in the couplings to generate specific mimics **9ai-hi** that represent controls for testing which contain only one mimic, triazine with the biotin label, and the morpholine part which represents the linker. Bivalent compound **9ii** featuring two morpholine moieties and the DTAB linker was also made as a complete control.

Our group has made some fluorescent or triethylene glycol derivatized tag labeled libraries in previous research.[43] Different from the established methods, DMSO was used as the solvent instead of THF or acetonitrile in the first monomer addition and a lower concentration was necessary due to reduced solubility of the Biotin-containing linker DTAB.

Scheme 3.2. Solution phase method for the preparation of the biotin-labeled library of bivalent compounds **9**.



Scheme 3.2. continued.



Thus a library of 45 biotin-labeled bivalent mimics was prepared using the method outlined in Scheme 3.2. Firstly, the monovalent compounds **4c-d**, **5c-d**, **6a-b** and **7a-b** were deprotected with TFA in absence of scavenger. The resulting unmasked intermediates dissolved in DMSO were reacted with equimolar amounts of DTAB and excess K_2CO_3 (4 equiv.) for 2-3 h to give the intermediates **8**. These intermediates in DMSO solution were split into small portions, and then coupled with another equivalent of the deprotected monomers **4c-d**, **5c-d**, **6a-b** and **7a-b** to give the desired bivalent compounds **9**. Alternatively, mophorline was used instead of the deprotected monomers as a capping group to generate a set of control compounds **9ai-ii**.

In this process, the purities of intermediates **8** were monitored by analytical HPLC, and they were used to judge whether the first coupling was complete and whether the purity itself was high enough to carry on the next step without purification. Fortunately, the reactions all completed in a few hours and met our purity criteria. After the second addition, the purities of bivalent compounds **9** were evaluated again by analytical HPLC (Figure 3.2a). Ideally, they would be more than 85 % and could be used for biological studies directly after a simple work-up and desalting procedure. However, we found that about 18 out of 45 bivalent compounds were satisfactory (≥ 85 %), 11 were not very pure (70-85 %), and another 1/3 of the molecules were less than 70 % pure, which definitely cannot be used directly for assays.

The impurities might be caused by several reasons. The major source of error was in weighing the compounds. Very slight differences in masses can cause serious effects in these mg-scale reactions. The oily monomers were much harder to weigh accurately in comparison with solid compounds. Another possibility is from pipetting errors. This has been found and proved in our previous work and could be corrected by simply repeating the syntheses for more times. Unfortunately for this particular case, there were not enough monomers to re-prepare the bivalent compounds to increase the purity. Other possible reasons for lower purities might include incomplete first addition and side-chain effects. Since most of the monomers are not very UV active at our observation wavelength (254 nm), it was hard to tell whether there were unreacted monomers left in the first step and this could cause byproducts in the second step. Amino acid side-chains that feature different functionalities might have a subtle influence on the monomer reactivity and reaction condition. However this did not cause much trouble in previous similar experiments according to our experience.

Finally we chose to purify all the bivalent compounds **9** (including the ones having a purity > 85 %) via RP-preparative HPLC. This surely provided us compounds pure enough for both binding and cellular assays. Most of the inorganic salt (K_2CO_3) was converted into KCl salt first by treating the lyophilized reaction residue with small amount of dilute HCl. After removing the solvent, the product was separated from the salt by extracting them into methanol; these were then concentrated and re-dissolved in appropriate solution for HPLC purification.

HPLC and MALDI analyses were performed on all the bivalent compounds **9** after purification. UV detection was used to determine the purities shown in Figure 3.2b. Purity data before the purification is given in Figure 3.2a. In summary, this dimerization methodology allows us to obtain a biotin-labeled bivalent compound **9** library from the monovalent compounds to combine with each other relatively cleanly in a one-compound-per-well format.

a.

	i	a	b	c	d	e	f	g	h
i(cap)	100								
a	100	75							
b	98	71	89						
c	100	72	68	80					
d	100	47	80	58	89				
e	95	45	59	43	71	64			
f	100	88	87	70	79	56	100		
g	98	93	92	70	58	41	72	86	
h	100	51	57	58	81	65	76	36	85

UV purities (%)

below 70

70 – 85

85 and up

b.

	i	a	b	c	d	e	f	g	h
i(cap)	100								
a	100	100							
b	100	98	100						
c	100	100	100	100					
d	100	100	100	100	100				
e	100	100	100	100	100	100			
f	100	100	100	100	100	100	100		
g	100	100	100	100	100	100	100	100	
h	100	100	100	100	97	100	100	100	100

UV purities (%)

below 85

85 – 90

90 and up

Figure 3.2. Purities of the library of compounds **9** where the UV detection method was set at 254 nm. **a.** crude purity; and **b.** after purification. The term 'cap' is used for morpholine.

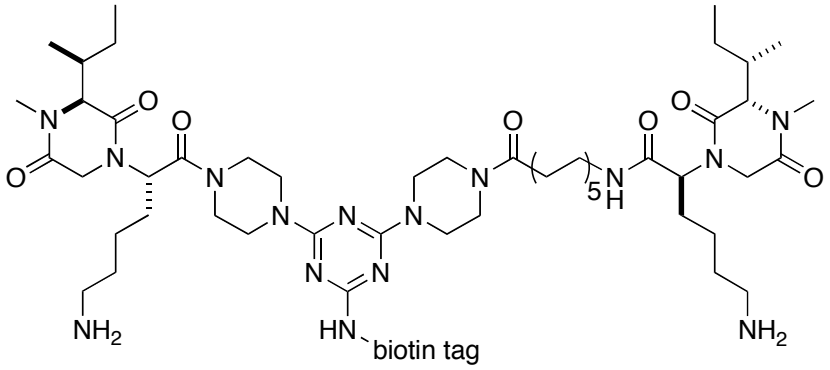
Twenty percent of the bivalent compounds **9** were characterized by one-dimensional and two-dimensional NMR experiments. (See Appendix B) All the compounds gave satisfactory molecular ions in MALDI-MS analyses. The product yields were not determined due to the small scale of the synthesis.

3.4 Biological Assays

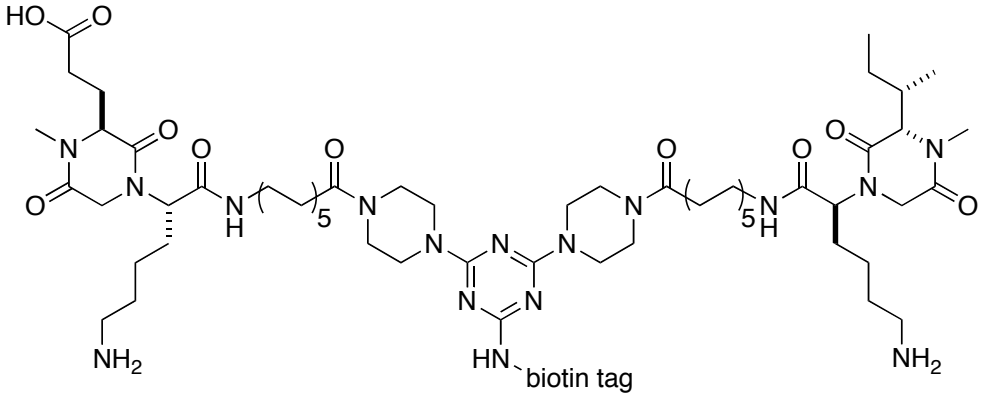
Binding assays were performed to assess the ability of peptidomimetics **9** to bind Trk and p75 receptors;[112,69,58,113] these were conducted by our collaborator, Dr. Uri Saragovi and his coworkers at McGill University in Canada.

The library of 45 bivalent compounds was tested for binding to transfected TrkC-expressing, TrkA-expressing, or p75-expressing cells in a fluorescence activated cell sorting (FACScan) assay. Live cells expressing the indicated receptors were incubated with ligands (50 μM) then with FITC-avidin, washed, and immediately studied by FACScan. Data was acquired and analyzed via CellQuest software. The NIH-IGF-1R cells which do not express any neurotrophin receptor served as a negative control.

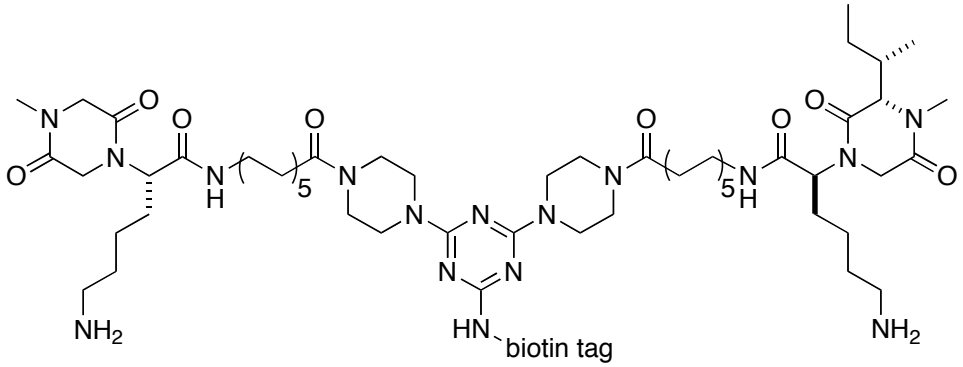
From the preliminary screening data, four bivalent compounds exhibited relatively high binding affinity and better selectivity against TrkC-expressing cells. The structures of the four compounds are illustrated in Figure 3.3. As shown in Figure 3.4, compounds **9gc**, **9hg, 9ge** and **9gg** bind to TrkC but not significantly to TrkA. Notably, the four "hits" all contain at least one cyclo(Ile-Lys)-long linker (*i.e.* **5c**) monovalent fragment. However, residues "IK", "EK" and "GK" are found in NGF turn sequences from different sources but not in NT-3, while residue "TG" corresponds to one of the turn sequences in NT-3 but none of the cyclo(Thr-Gly) mimics showed high affinity in binding assay probably due to their relatively poor water solubility.



9gc

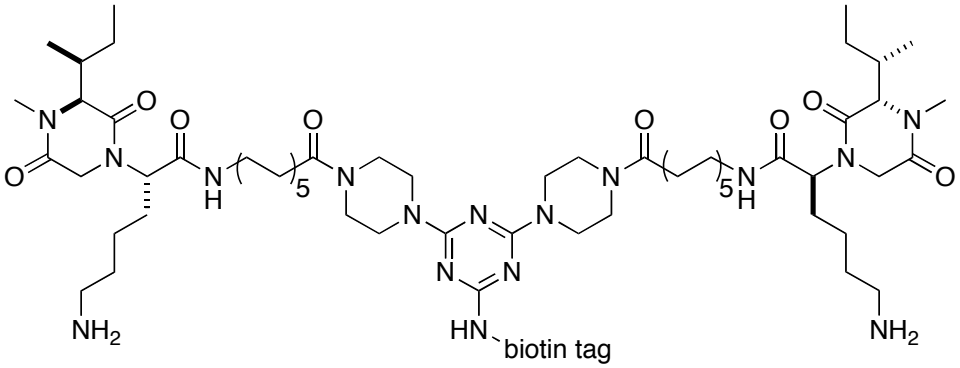


9hg



9ge

Figure 3.3. Structures of compounds **9** that bind to TrkC from the preliminary FACScan assay.



9gg

biotin tag =

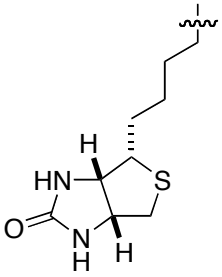


Figure 3.3. continued.

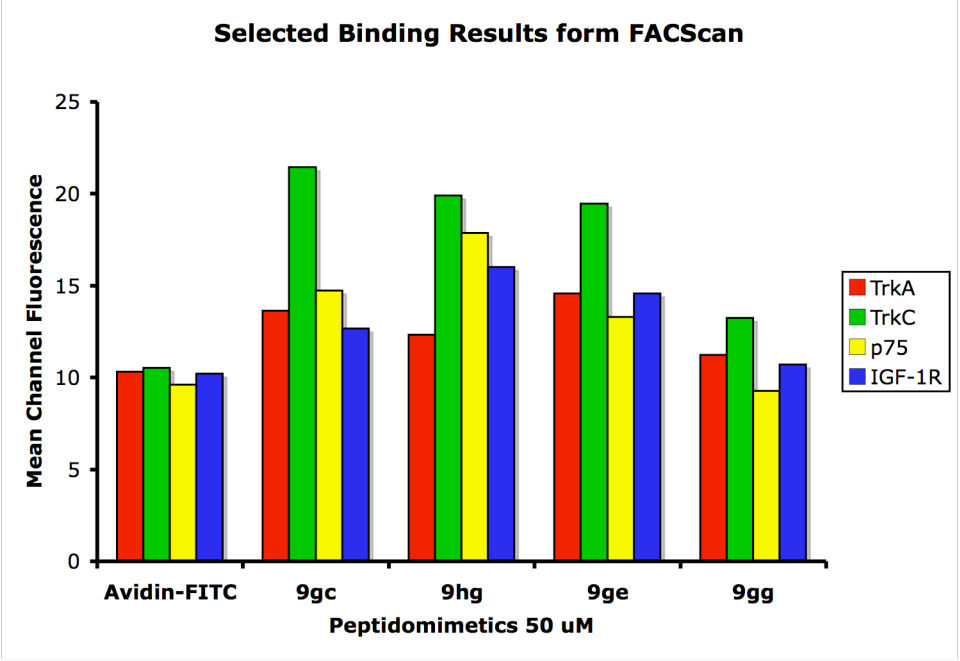
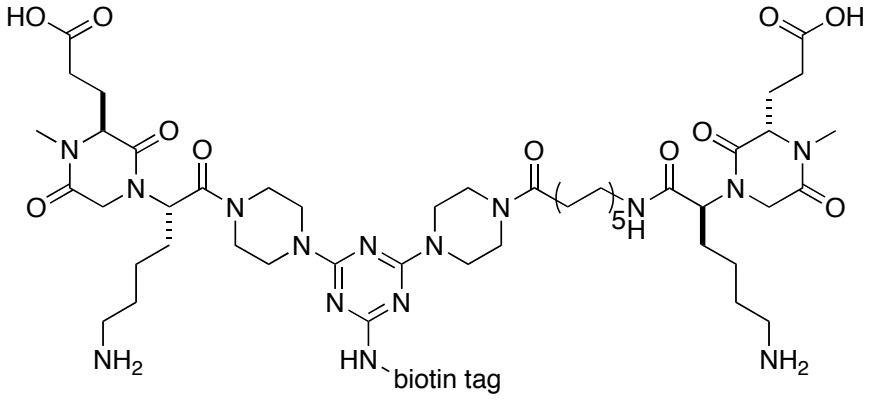
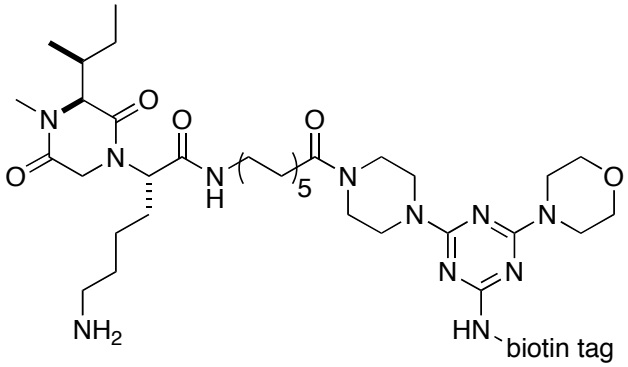


Figure 3.4. FACscan assay data and the graph presentation of the four "hits".

In subsequent cell survival assays, three bivalent compounds **9hd, 9gi** and **9gg** were found to block the trophic activity of NT-3 in TrkC cells at 10 uM, and no significant effect was observed for these compounds in IGF-1R and TrkA cells. (Figure 3.5 and 3.6) The survival studies on these compounds will be repeated with different doses to confirm their selectivity for TrkC cells. Among the three, compound **9gg** is worthy of attention for its preferential binding to TrkC receptors shown in previous FACScan assays, while the other two mimics did not show the same trend. More interestingly, **9gg** is a homodimer of two cyclo(Ile-Lys)-long linker moieties. Overall, there is no significant correlation found among these active compounds.



9hd



9gi

Figure 3.5. Structures of compounds **9** that block the trophic activity of NT-3 in TrkC cells in cell survival assay (**9gg** is shown in Figure 3.3).

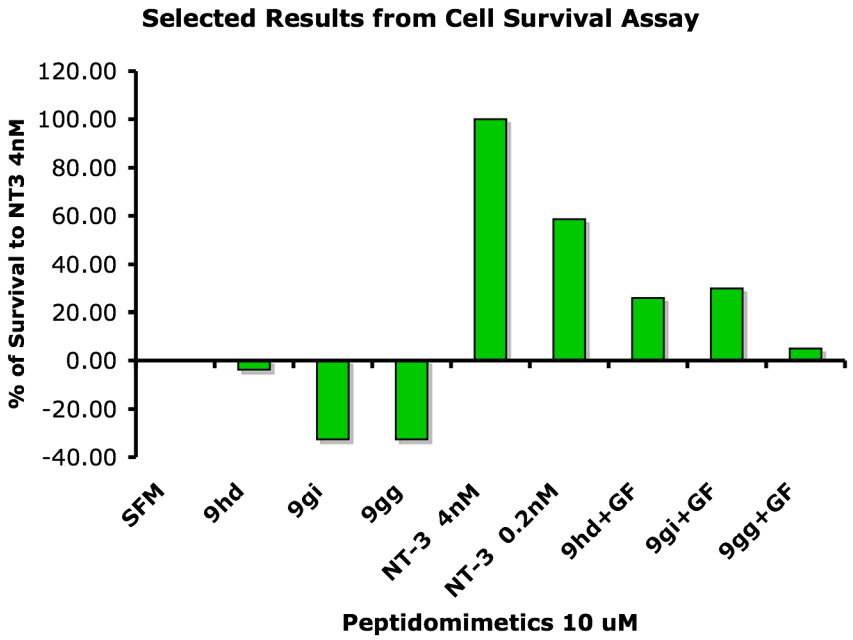


Figure 3.6. Cell survival assay data and the graph presentation of the three "hits".

3.5 Summary

The diketopiperazine (DKP) scaffold has proven to be a versatile template in combinatorial chemistry for its 4 ring atoms that are basis for the generation of molecular diversity.[80] Their chiral, rigid and functionalized structure is favored for binding to various receptors with high affinity and gives a broad range of biological activities.[79]

A library of eight monovalent DKP mimetics **4c-d**, **5c-d**, **6a-b** and **7a-b** is prepared from corresponding dipeptides via intramolecular S_N2 cyclization reactions in solution. These DKP compounds contain two amino acid side-chain functionalities that are both derived from natural amino acids. Parallel solution phase synthesis allows quick and efficient generation of the desired peptidomimetics in good yields. Each monomer was attached to a short or long linker before they were combined into dimers.

The eight DKP mimics (along with morpholine as 'capping' group) were assembled into a library of 45 bivalent molecules **9** via an efficient combinatorial strategy previously reported by our group. The two monovalent mimics were coupled to the

linker scaffold sequentially in solution. A biotin 'tag' was attached to each bivalent molecule to facilitate further binding and biological assays. The purities of bivalent compounds **9** were evaluated by analytical HPLC before and after purification.

The fluorescence-activated cell sorting (FACScan) assay has performed on the biotin-labeled compounds for binding to transfected TrkC-expressing, TrkA-expressing or p75-expressing cells. Four bivalent compounds **9gc**, **9hg, 9ge** and **9gg** have exhibited relatively high binding affinity and better selectivity against TrkC-expressing cells at 50 uM. All of these molecules contain at least one cyclo(Ile-Lys)-long linker fragment. In cell survival assay, three bivalent compounds **9hd, 9gi** and **9gg** blocked the trophic activity of NT-3 in TrkC cells at 10 uM, and no significant effect was observed for these compounds in IGF-1R and TrkA cells. The survival on these compounds will be repeated with doses to confirm their selectivity for TrkC cells. Interestingly, compound **9gg,** which also showed preferential binding to TrkC receptors, is a homobivalent peptidomimetic of two cyclo(Ile-Lys)-long linker moieties.

CHAPTER IV

SOLUTION-PHASE SYNTHESIS OF OXADIAZOLE-BASED BETA-TURN MIMETICS AND ATTEMPTS TOWARDS BIVALENT MOLECULES

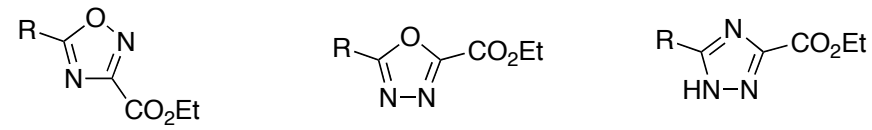
4.1 Introduction

Screening of small molecule is the most viable way for the identification and optimization of potential bioactive compounds. The main goal of our approach is to mimic the complex molecular interactions of natural proteins by designing small molecules that mimic protein secondary structures, especially β-turns. This allows for the functionality identification of important 3D conformations from a molecular perspective and helps further the development of therapeutic agents. Usually the unique conformational alignment of functional groups of proteins is relvent to their molecular recognition. Therefore peptidomimetic probes can obtain rich, useful information from structure-activity relationship studies and conformational analyses of bioactive peptides and proteins. An ideal solution is to combine selective and metabolically stable analogs, peptide backbone modifications and relatively constrained amino acids together into bioactive peptide sequences.

5- and 6-membered heterocyclc moieties are common structure motifs in bioactive compounds, such as natural products and synthetic molecules, which play enormous impact on their bioactivity.[114-120] These heterocycles rigidify otherwise flexible backbones, contribute to lipophilicity, and their presence frequently correlates with protein binding properties. Thus the alternation and variation of these heterocycles are of major interest for the derivatization of these bioactive molecules. Moreover, their small and rigid, conjugate and stable structures make them ideal scaffolds for mimicry of various protein motifs.

For example, there are reports on the uses of hetercyclic rings in peptidomimetics to replace peptide fragments to convert peptides into chemically stable and orally available molecules (Figure 4.1).[37] There are also many examples in the literature of

using heterocycle derivatives as different types of protein recognition motif mimics (Figure 4.2).[39,121]



R = amino acid side-chains from Ala, Gly, Asp, Phe, Ser, Arg, Cys, Pro

Figure 4.1. Three heterocyclic ring systems as Xaa-Gly dipeptide mimetics.

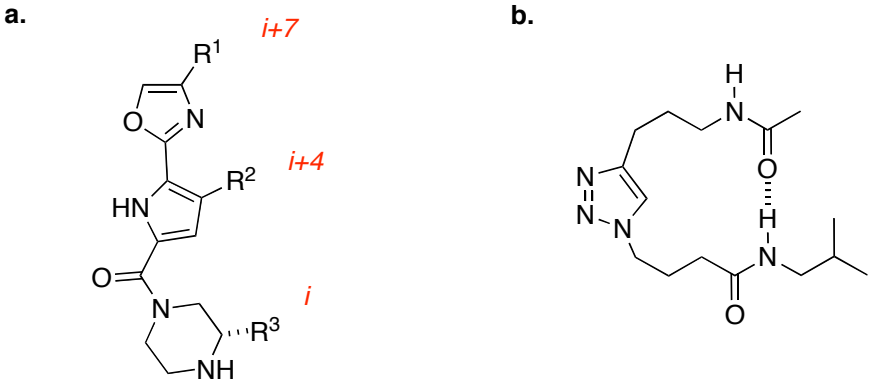


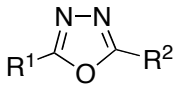
Figure 4.2. Examples of different azole type compounds as peptidomimetics. **a.** non-peptidic tricyclic oxazole-pyrrole-piperazine scafold as α-helix mimetic; **b.** triazole ring-based new β-turn mimic.

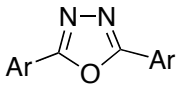
Our group has been focusing on the design and synthesis of small molecules that mimic or disrupt protein-protein interactions and have gained some experience in designing β-turn mimics.[43,58] In this work we design and synthesize a series of novel β-turn mimics. To mimic most possible protein motifs we use a heterocycle-based scaffold to obtain necessary backbone rigidity and incorporate the natural amino acid side-chains' functionalities in somewhat restrained conformations. These amino acids were chosen

from the most commonly found bioactive sequences in various protein-protein interactions. Therefore these designated peptidomimetics bear peptidic and non-peptidic characteristics at the same time. Most of these compounds comply with Lipinski's 'rule-of five' to be more 'drug like'.[122]

It is always important to start from an appropriate design to accquire good small molecule mimetics. From the screening of many heterocyclic building blocks, we found that 1,3,4-oxadizole is a class of compounds that is thermally stable neutral and fully conjugated.[123] Their chemistry has been known for more than 100 years.[124,125] They have broad uses in many areas, for example in drug synthesis, scientilation materials, polymer and dyestuff industry.[126,127]

Many oxadizole containing compounds have been reported to present various bioactivities. In Figure 4.3, bactericidal and/or fungicidal activity was reported for oxadiazoles **(a)** and amino-oxadizoles **(b)**.[128] Analgesic activity, antiinflammatory and antiproteolytic properties, anticonvulsant and nervous system depressant activity, and local anaesthetic activity was shown by amino-oxadiazoles **(c)-(f)**, respectively.[129-131] Diaryloxadiazoles **(g)** exhibit antiinflammatory, sedative and analgesic properties.[132]





(a) R^1 = Ar, R^2 = $CH_2CONHCONHR$
(b) R^1 = $ArOCH_2$, R^2 = NHCOR
(c) R^1 = 2-pyridyl, R^2 = NR_2•HCl
(d) R^1 = 4-biphenylyloxymethyl, R^2 = NHAr
(e) R^1 = Ar, R^2 = $NHCH_2CONHR$
(f) R^1 = $ArOCH_2$, R^2 = $NHCO(CH_2)NRR'$•HCl

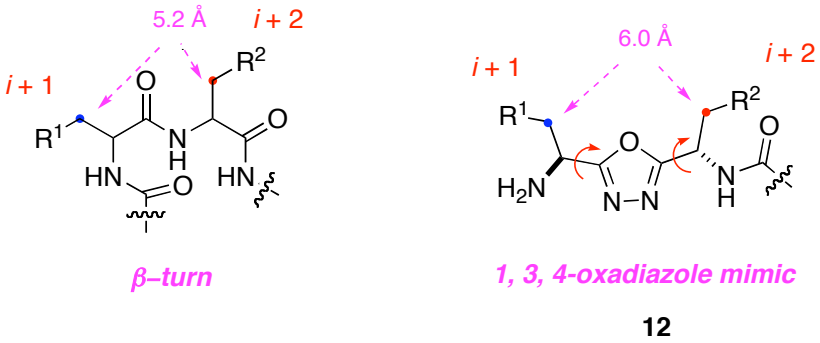
(g) Ar = trimethoxy- or 3,4-dioxymethylene-phenyl

Figure 4.3. Examples of bioactive 1,3,4-oxadiazoles.

Overall they are interesting synthetic targets and are ideal for our small molecule design as potencial turn analogs; this is further illustrated by the result of our molecular modeling study (Figure 4.4). While reintaining the side-chain functionalities of natural

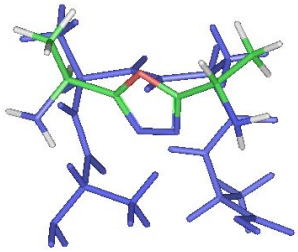
amino acids, our efforts to develop small molecule mimics **12** with less peptide character include using a 2,5-substituted 1,3,4-oxadizole as the backbone to bridge two amino acid side-chains corresponding to the ones found in the loop region of the neurotrophins. The big advantage of this design is that the target compounds can be relatively easily made on large scale, and the natural amino acids that correspond to original protein sequences are highly accessible. Conceptually the side-chains at R^1 and R^2 can mimic $i + 1$ and $i + 2$ positions of a standard β-turn structure by adjusting two degrees of freedom.

a.



12

b.



overlay with type I β–turn

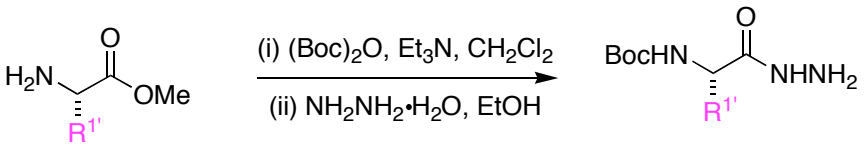
Figure 4.4. **a.** The key distance of C^b-separations of the $i + 1$ to $i + 2$ residues of a type I β-turn and of the monovalent turn mimics **12**; **b.** an overlay of the 1,3,4-oxadiazole mimic onto a type I β-turn.

4.2 Preparation of 1,3,4-Oxadiazole Based Beta-Turn Mimetics

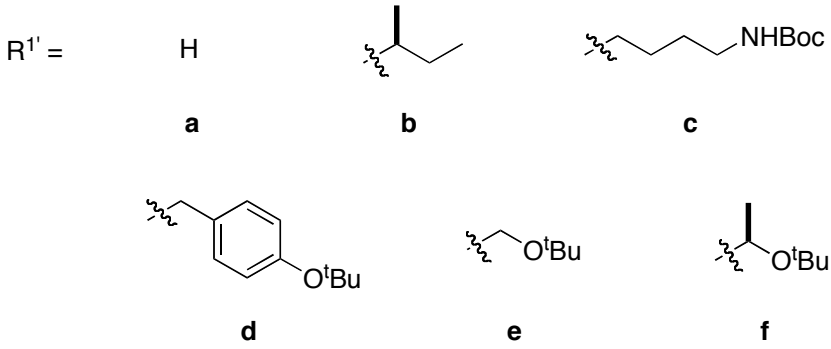
In this work, a library of twenty 1,3,4-Oxadiazole mimetics was prepared in solution. At first a set of Boc-protected amino acid methyl esters was converted to acid hydrazides[133] by reacting with hydrazine monohydrate in alcohol. The Boc-protected hydrazides **9a-f** were then coupled with different Cbz-protected amino acids[134] to produce the diacid hydrazides with more functionality. Using the diacid hydrazide as a cyclization precursor, the cyclodehydration[19,135,136] proceeded cleanly with activated triphenylphosphine to offer the desired 5-member ring. The Cbz protective group of seventeen synthesized 1,3,4-oxadiazole compounds **11a-q** was removed under standard hydrogenolysis conditions. Boc-Inp-OH as the short linker for this library was coupled to the cyclic core by forming an amide bond to give monovalent mimics **12a-q** in good yields.

Scheme 4.1. General synthetic scheme of 1,3,4-oxadizole mimetics. **a.** synthesis of amino acid hydrazide fragments; **b.** synthesis of 1,3,4-oxadizole intermediates **11a-q**; **c.** synthesis of monovalent oxadiazole mimics **12a-q**.

a.



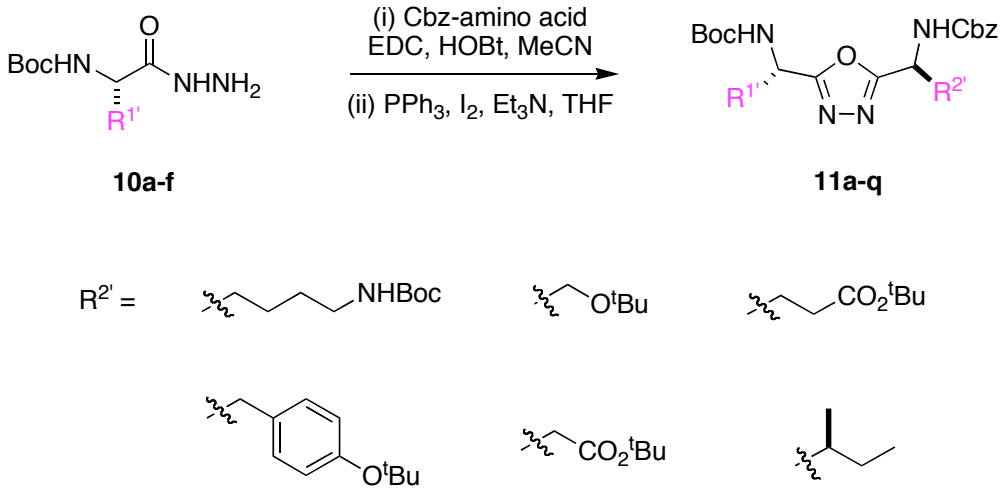
10a-f



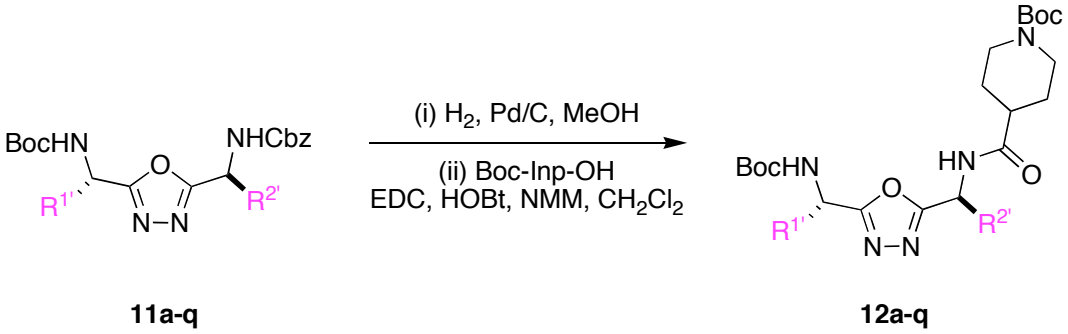
a **b** **c** **d** **e** **f**

Scheme 4.1. continued.

b.



c.



The library of twenty monovalent peptidomimetics **12a-t** is listed below (Table 4.1). All the compounds were functionalized with Boc-Inp-OH moiety as a short linker. Seventeen of them were prepared following the route shown in Scheme 4.1 while three (**12r-t**) were prepared from *N*,*N*-dimethyl substituted Serine methyl ester instead of Boc-protected Serine methyl ester (Scheme 4.2). Reductive alkylation provided *N*,*N*-dimethyl serine methyl ester **13**,[137,138] which was transformed to the corresponding acid hydrazides **10g**. The remaining synthesis steps were carried through following the above general synthetic scheme to offer three more oxadiazole compounds **12r-t** with a short linker through intermediates **11r-t**.

Scheme 4.2. Synthesis of *N*,*N*-dimethyl serine methyl ester **10g** and intermediates **11r-t**.

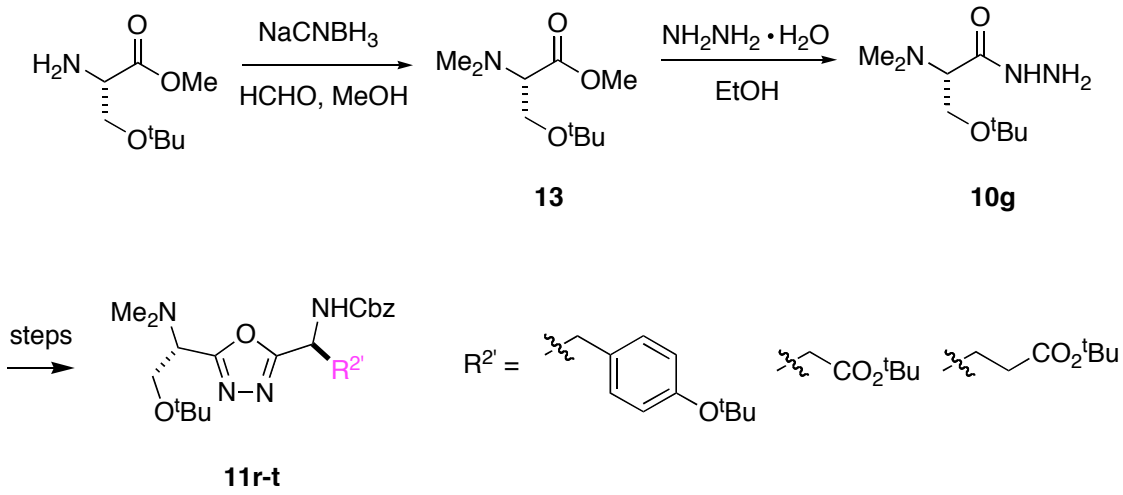
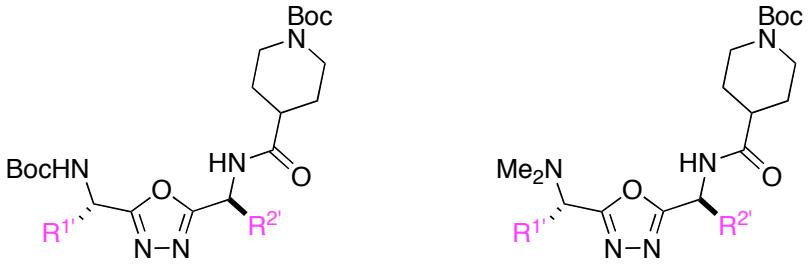


Table 4.1. Summary of 1,3,4-oxadiazole library **12a-t**.



12a-q **12r-t**

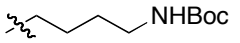
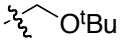
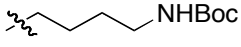
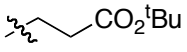
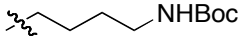
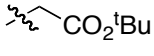
comp'd **12**	scaffold	R^1	R^2	yield (%)[a]
a	I			70
b	I			77
c	I			71

Table 4.1. continued.

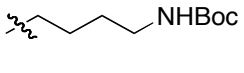
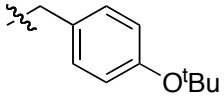
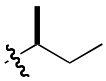
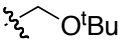
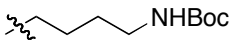
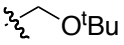
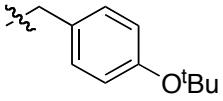
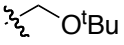
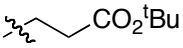
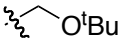
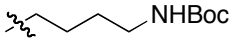
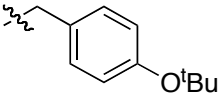
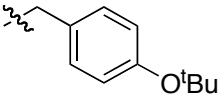
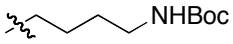
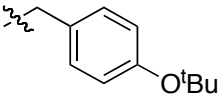
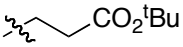
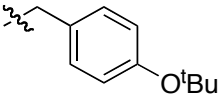
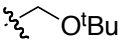
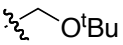
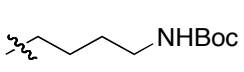
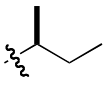
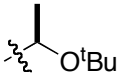
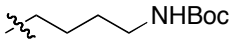
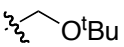
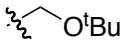
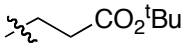
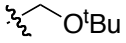
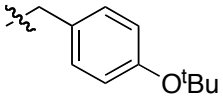
comp'd **12**	scaffold	R^1	R^2	yield (%)[a]
d	I			93
e	I			78
f	I	H		83
g	I			84
h	I			80
i	I			82
j	I			77
k	I			74
l	I			78
m	I			89
n	I			85
o	I	H		74
p	I			92
q	I			70
r	II			82

Table 4.1. continued.

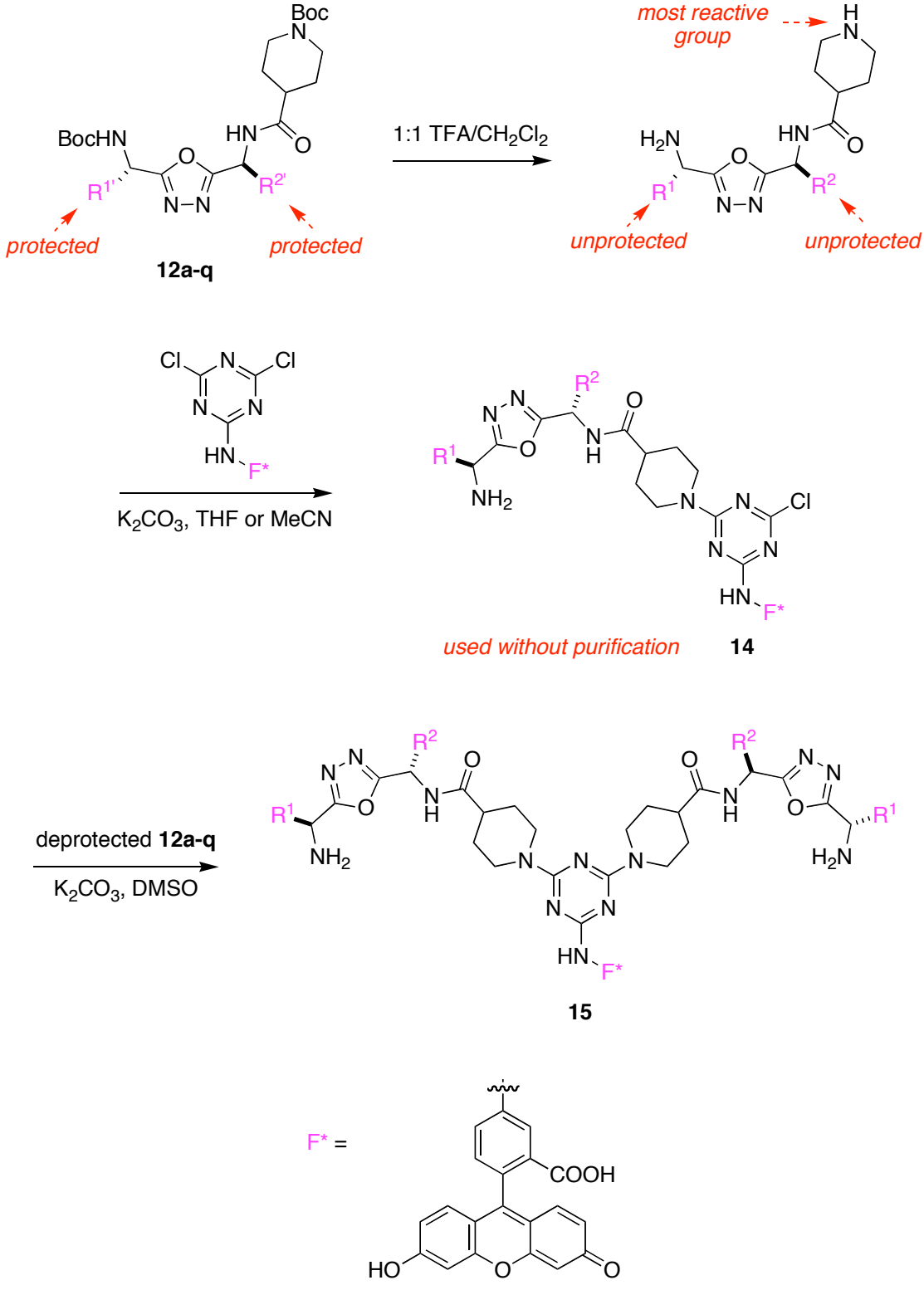
comp'd **12**	scaffold	R^1	R^2	yield (%)[a]
s	II			89
t	II			67

[a] Calculated on the basis of intermediate **11a-t**.

4.3 Preparation of Bivalent Compounds

The preparation of the bivalent library of compounds **15** is similar to that discussed in Chapter III.[43] As shown in Scheme 4.3, all the acid labile protecting groups of the monovalent compounds in Table 4.1 were removed by 50 % TFA and treated with DTAF (5-(4,6-dichloro-s-triazin-2-ylamino)fluorescein hydrochloride) and K_2CO_3 in THF for 4-6 h at 25 ^oC. The resulting intermediate **14** was concentrated and redissolved in DMSO and split into small portions. An equimolar amount of another deprotected monomer was added to each portion followed with the addition of excess of K_2CO_3. The resulting mixture was placed on a shaker and reacted for 72 h to afford bivalent compounds **15**.

Scheme 4.3. Preparation of bivalent mimetics **15**.



Purities of the crude materials of each step were evaluated by analytical HPLC, monitored at 254 nm by a UV detector. Usually the crude purity after first monomer attachment is more than 95 % in order to achieve a purity more than 85 % for the final bivalent compounds. However, for this library, the purities of intermediates **14** generated after the first addition are not good enough to carry forward with the dimer synthesis. Most of the compounds did not give satisfactory purities even after the first reaction step with DTAF. There were always some impurities in the crude product **14** from analytical HPLC traces. The same reactions were repeated several times and with different substrates and we still could not obtain satisfactory results. It seems that this method may not be ideal for these oxadizole compounds.

By analyzing the structure of the monomers **12a-q** and reaction conditions, I think there are several possible reasons that may cause the reactions to be not very clean. (i) deprotection of the monomer is not complete, (ii) deprotection generates some impurities, (iii) functional groups on amino acid side-chains interfere with the S_NAr reaction, (iv) the free amino group next to R^1 interferes with the S_NAr reaction.

Among these possibilities, I first ruled out the incomplete deprotection by checking NMR and HPLC after deprotection. Then I started to examine the last one through structure modification. The free amino group next to R^1 was replaced with fully substituted *N*,*N*-dimethyl group. As was previously explained (section 4.2), this was achieved by using a slightly different starting material. A total of three 1,3,4-oxadizole compounds **12r-t** were prepared starting from *N*,*N*-dimethyl serine methyl ester **10g**. The three compounds were tested under the same reaction conditions and compared with unsubstituted serine compounds **12a-q**. After the reaction was complete (no more DTAF peak detected), I found that some major impurity peaks disappeared such as those used to be found from the HPLC chromatograms of **12i** (SE mimic) and **12g** (SD mimic), while some did not change much (*i.e.* **12h**, a SY mimic). It seems that the free amino group does affect the S_NAr reaction. Because when this amino group gets too reactive it tends to compete with the NH from the short linker, which theoretically is most reactive group in the whole molecule. It will react as a nucleophile with DTAF and lead to a complex byproduct **16** as shown in Figure 4.5a as an example, which has been verified by LCMS

analysis. This also explains why the same impurity does not appear anymore when the specific amino group is blocked. It should be noted that this result depends on the amino acid side-chains in the system, as some of them did not have this problem (*e.g.* SY mimic **12h**).

However this did not fully solve the problem to obtain a clean reaction. From analytical HPLC there are still some detectable impurity peaks, although very small, and these are very hard to identify. I assume they are from the residues of acid cleavage (reason (ii) as mentioned above), however it is not easy to prove and this has not been done yet. The deprotected compounds are not UV active under 254 nm and are extremely polar so that they are not ideal for HPLC analysis.

a.

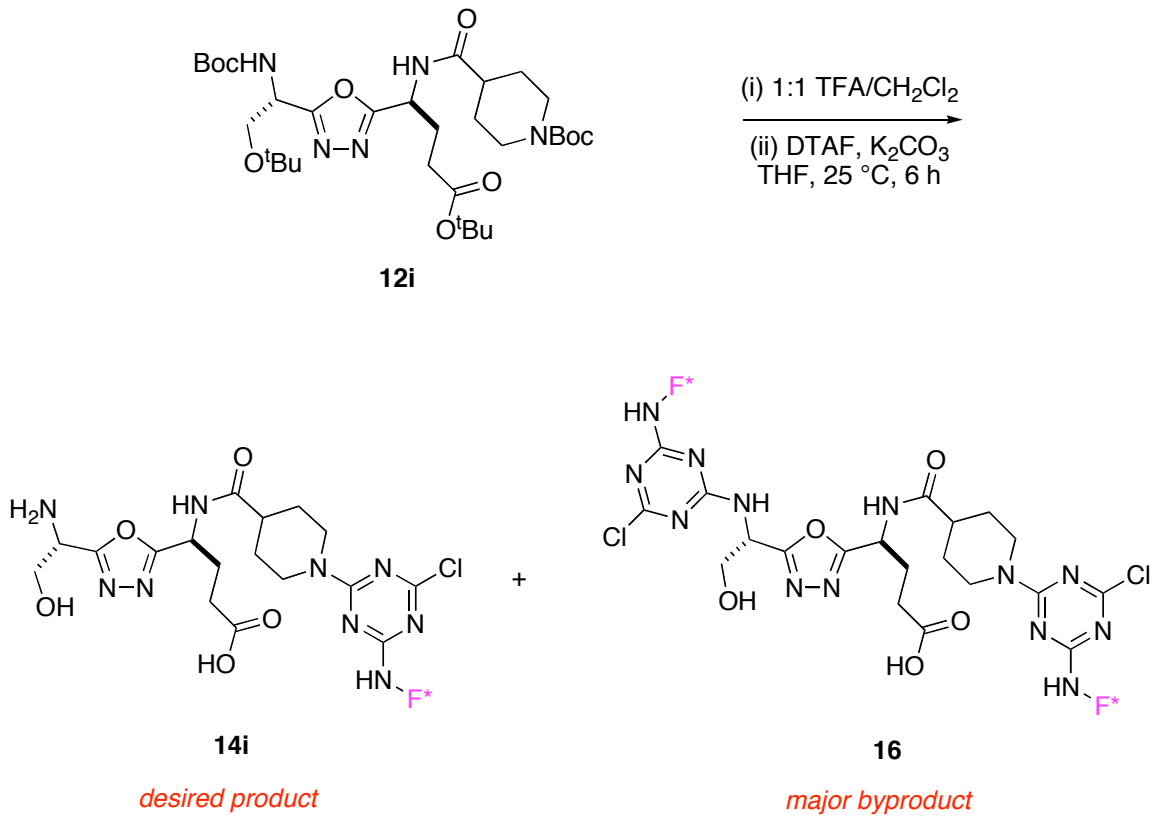
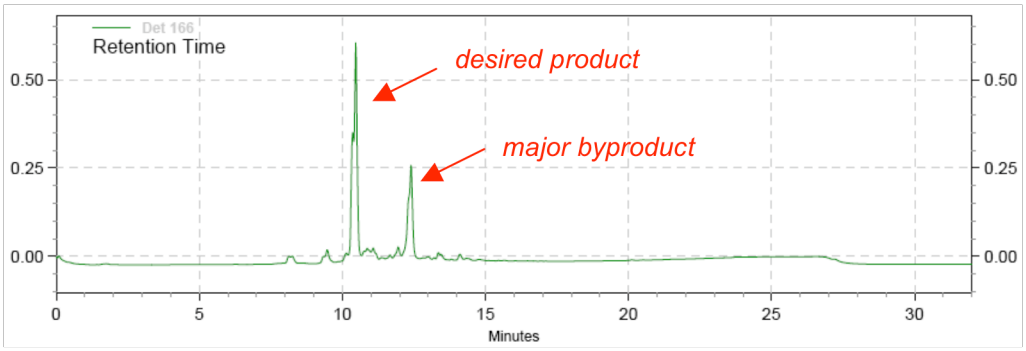
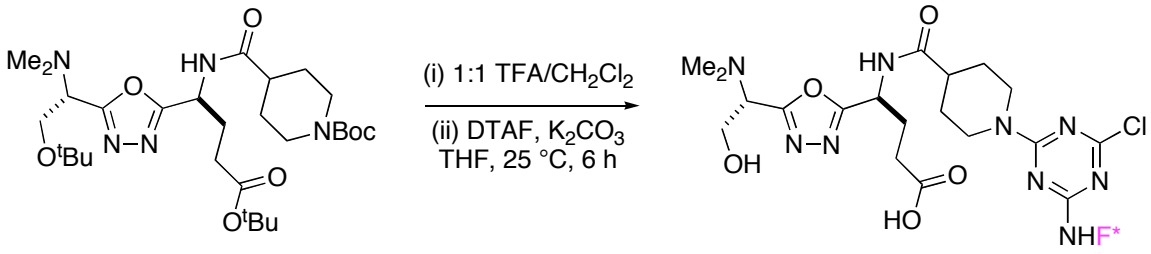


Figure 4.5. a. Reaction of monomer **12i** with DTAF (1:1) and the HPLC trace (UV at 254 nm) of its crude purity; **b.** reaction of monomer **12s** with DTAF (1:1) and the HPLC trace (UV at 254 nm) of its crude purity.



b.



12s

14s

desired product

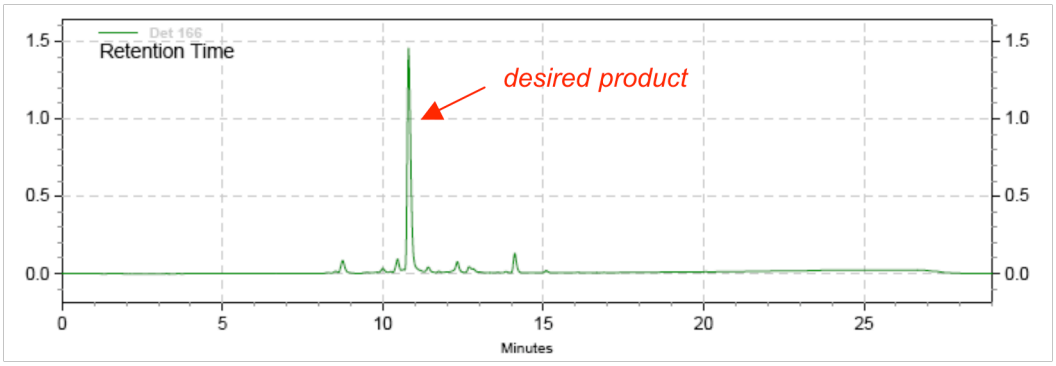


Figure 4.5. continued.

4.4 Biological Assays

Since the twenty 1,3,4-oxadiazole compounds **12a-t** were not made into a library of labeled bivalent compounds, they could not be tested in binding assays. However, we were still able to test them in cell survival assay for their biological functionalities. Our collaborator, Dr. Uri Saragovi and his coworkers at McGill University in Canada are still working on testing these compounds.

Forty 1,3,4-oxadiazole compounds including twenty **12a-t** in fully protected form and twenty **17a-t** in fully deprotected form (not shown) were sent to NIH screening centers for extensive biological assays. Several compounds were found to be active in different assays (these results and more details of the bioassays can be obtained from http://pubchem.ncbi.nlm.nih.gov).

4.4.1 Screen for Chemicals that Inhibit the RAM Network

This is a selection-based screening of compound libraries to discover highly specific small molecule inhibitors of the budding yeast RAM network. The yeast strain used in the screen was engineered to contain an "analog-sensitive" mutant allele of the RAM network protein kinase Cbk1, which is susceptible to inhibition by 1-napthyl PP1 (1-NA-PP1). This permitted pre-inhibition of URA3 expression, a step that was necessary to ensure that cells did not begin growth under screening conditions with lethal levels of this gene product, and also allowed the use of 1-NA-PP1 as a positive control drug. Percent inhibition of the RAM network signaling was calculated using the optical density in control wells with untreated cells as full RAM network activity (no growth; 0% inhibition) and wells treated with 10 μM 1-NA-PP1 as complete inhibition of the RAM network (full growth; 100% inhibition). Library compounds were screened at 50 μM.

Compounds in this primary screen were scored on a scale of 0-40 based on the lowest activity from the two runs where 0 corresponds to no inhibitory activity and a score of 40 corresponds to 100% inhibition. An activity score of 4 or higher corresponds to compounds with an active outcome.

In this screening assay for chemicals that inhibit the RAM network conducted at southern research molecular libraries screening center, two compounds **12i** (SE mimic) and **12c** (KD mimic) showed some activity and scored 12 and 11, respectively.

4.4.2 Primary Cell-based High Throughput Screening Assay for Inhibitors of Wee1 Degradation

Previous studies suggest that nuclear tyrosine kinase Wee1 may act as a tumor suppressor. The identification of probes that selectively increase levels of Wee1 may provide a potent route for cell cycle-based cancer therapy. Compounds that inhibit Wee1 degradation will provide information about the role of Wee1 in mitotic entry and cell cycle progression. The hypothesis is based on studies that show that Wee1 stabilization inhibits mitotic entry. Moreover, the liabilities of several reported intriguing compounds necessitate the discovery of high affinity, selective probes for Wee1 stabilization. The assay uses HeLa cells transfected with a kinase negative mutant of Wee1 (Wee1K328M) fused to a luciferase reporter gene. As designed, compounds that increase Wee1K328M-luciferase stability and/or prevent its degradation will lead to increased well luminescence compared to untreated wells. Specifically, compounds that increase luminescence are considered Wee1 degradation inhibitors.

The reported activity score has been normalized to 100% of the highest observed activation. The inactive compounds of this assay have activity score range of 0 to 6 and active compounds range of activity score is 6 to 100.

From the data provided by the Scripps research institute molecular screening center, the activity ranking score of compound **12a** (KS mimic) and **12s** (S'E mimic) were 8 and 7, which are slightly higher than 6 to be in the active range.

4.4.3 μHTS Identification of Diaphorase Inhibitors and Chemical Oxidizers: Counter Screen for Diaphorase-based Primary Assays

The purpose of this assay is to identify inhibitors of diaphorase, which is an enzyme that reversibly catalyzes the reaction of converting NAD(P)+ to NAD(P)H and transfers its electrons to a variety of Redox dyes. It is expected that strong inhibitors of diaphorase would appear as inhibitors in the PMI assay. Thus, diaphorase is a counter-screen for the PMI and other diaphorase-based assays allowing identification of the artifacts of the coupled assay detection system. In addition, compounds that have strong oxidative properties are also expected to suppress the resorufin formation and appear as inhibitors of diaphorase based reactions. Thus, the assay could be utilized for library

characterization in respect to redox-active compounds. Compounds with greater than 50% inhibition at 20 μM concentration are defined as actives of the primary screening, and the primary screening actives proceed to the dose-response confirmation stage. However, no dose-response data has been published on PubChem so far.

First tier (0-40 range) of the scoring system is reserved for primary screening data. The score is correlated with % displacement in the assay demonstrated by a compound at 20 μM concentration.

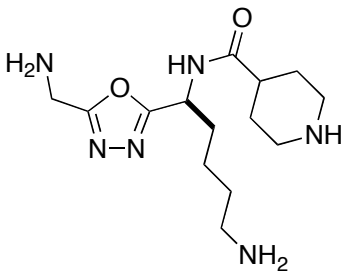
Compound **12j** (SK mimic) scored 34 in this assay based on the result from San Diego center for chemical genomics.

4.4.4 μHTS of Mcl-1/Bid interaction inhibitors

Currently there is still a lack of small-molecule inhibitors of Mcl-1. Reports indicate that the most potent inhibitor of the anti-apoptotic proteins Bcl-2, Bcl-xL and Bcl-w described to date, fail to bind to Mcl-1. This is evidence that Mcl-1 has critical structural differences compared with other anti-apoptotic proteins and strongly suggests that specific targeting of Mcl-1 should be possible. Bid is a broad acting Bcl-2 homology domain 3 (BH-3) only protein that binds Mcl-1 as well as other pro-survival proteins, including Bcl-2 and Bcl-xL. To screen for small molecular inhibitors that disrupt the interaction of Mcl-1 protein and its binding partners, a homogeneous time-resolved fluorescence resonance energy transfer (TR-FRET) assay was developed. In this assay, interaction of Mcl-1 protein with Bid peptide brings two conjugated fluorophores into proximity, leading to an energy transfer from europium to Dy647 and the generation of FRET signals.

This assay is used for the screening of the NIH/DPI library for Mcl-1/Bid interaction inhibitors. Compounds that cause > 50% inhibition are defined as positives. The percent inhibitions are rounded to 0 decimal places. Percent inhibitions > 100 are rounded to 100, and those < 0 are rounded to 0.

Compound **17j** (derivative of **11j** as a GK mimic) scored 90 in this assay conducted at Emory university molecular libraries screening center. It is also the only unmasked 1,3,4-oxadizole compound that was found active in these bioassays so far.



17j

4.5 Summary

Syntheses of rationally designed peptidomimics normally involve peptide bond surrogates or scaffolds wherein the critical side-chains are retained in a constrained conformation to maintain the desired biological properties. Our efforts toward the synthesis of small molecule β-turn mimetics used a 1,3,4-oxadizole based scaffold and two amino acid side-chains to mimic the distance and orientation at *i*+1 and *i*+2 position. The natural amino acid side-chains can be sequentially incorporated via amide coupling and stabilized via ring cyclization reactions.

This method produced a library of twenty β-turn mimics **12a-t**. These compounds contain a variety of amino acid side-chains, *ca*. glycine, isoleucine, lysine, serine, glutamic acid, aspartic acid, theronine, and tyrosine. A short linker was attached to the monomers via an amide bond. After removal of all the protecting groups, the monovalent compounds were introduced onto a fluorescent-labeled triazine scaffold sequentially. Unfortunately for some reason the reaction did not proceed cleanly. Studies have been carried out and some problems could be partially solved. More efforts are still needed to obtain a bivalent compound library.

The biological activities of these monomers are still under investigation. So far, based on the results from NIH screening centers, there are six compounds that showed certain activities in 4 bioassays of different types.

CHAPTER V

SYNTHESIS OF CYCLIC PEPTIDOMIMETICS OF PROTEIN A AND SMALL MOLECULE MIMETICS OF PROTEIN G FOR PURIFICATION OF IMMUNOGLOBLIN G VIA AFFINITY CHROMATOGRAPHY

5.1 Introduction

An antibody is defined as "an immunoglobulin capable of specific combination with the antigen that caused its production in a susceptible animal." Antibodies, which are produced in response to the invasion of foreign molecules in the body, exist as one or more copies of a Y-shaped unit, composed of four polypeptide chains. Each Y-shaped unit contains two identical copies of a large heavy chain, and two identical copies of a small light chain, named as such by their relative molecular weights. Antibodies can be divided into five classes: IgG, IgM, IgA, IgD and IgE, based on the number of Y units and the type of heavy chain. It is also the heavy chain that determines the subclass of each antibody, and each class of antibody has different abilities to find and help remove antigen in the body.

The most commonly used antibody in immunochemical procedures is of the IgG class because they are the major immunoglobulin (Ig) released in serum. For instance, it constitutes about 75 % of all Ig's found in serum. The classical Y shape of IgG is composed of the two variable, antigen specific F(ab) arms, which are critical for actual antigen binding, and the constant Fc tail that binds immune cell Fc receptors and also serves as a useful "handle" for manipulating the antibody during most immunochemical procedures (Figure 5.1). The number of F(ab) regions on the antibody corresponds with its subclass and determines the valency of the antibody. Direct-conjugated antibodies are labeled with an enzyme or fluorophore in the Fc region. The Fc region anchors the antibody to the plate in Enzyme-Linked ImmunoSorbent Assay (ELISA) procedures and is also seen by secondary antibodies in immunoprecipitation, immunoblots and immunohistochemistry. These three regions can be cleaved into two F(ab) and one Fc fragments by the proteolytic enzyme Papain, or into just two parts: one $F(ab)_2$ and one Fc

at the hinge region by the proteolytic enzyme Pepsin. Fragmenting IgG antibodies is sometimes useful because F(ab) fragments will not precipitate the antigen; and they will not be bound by immune cells in live studies because of the lack of an Fc region. Often, because of their smaller size and lack of crosslinking (due to loss of the Fc region), Fab fragments are radiolabelled for use in functional studies. Interestingly, the Fc fragments are often used as blocking agents in histochemical staining.

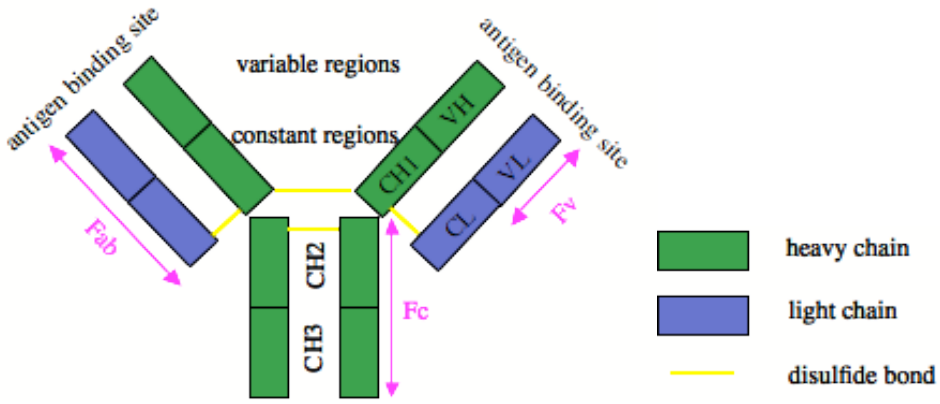


Figure 5.1. Schematic view of Y-shaped IgG.

IgG has widespread and varied applications, including ones in diagnostics,[139,140] as bioaffinity ligands in purification of high-value pharmaceuticals (*eg* cytokines and blood-clotting factors)[141] and as probes in diverse biochemical experiments.[140,142-145] Perhaps the most exciting potential applications of IgG are in the area of therapeutics.[146,147] Animal IgG can be "humanized" so that it is not rejected by the human immune system.[148] Consequently, time- and cost-effective techniques for large-scale isolation and purification of IgG are extremely important.

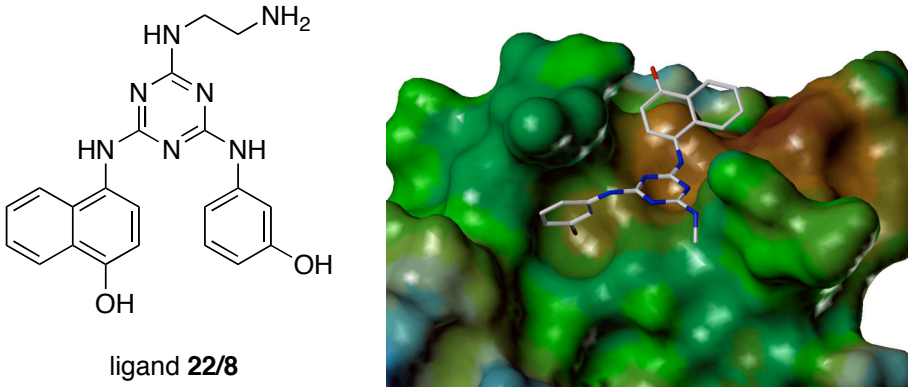
IgG is obtained from human plasma, ascites fluid, fetal calf serum culture supernatants of engineered cells (*eg* hybridoma and bacterial cells) and other sources, then it must be purified for most applications. However, the purity of IgGs is critical for therapeutic applications. Usually this is achieved via a chromatographic technique, and there are several options for this. Size-exclusion and ion-exchange[149] [150]methods have

both been used, but affinity chromatography is the most widely applied.[141,151,152] Monoclonal antibodies (mAbs) designed to bind a specific antigen may be purified using supports based on that substance.[153] A drawback of this approach is that the binding constants for the antigen tend to be so high that harsh elution conditions are required and denaturation of the IgGs can occur. Moreover, for widespread large scale purification of mAbs, methods that work irrespective of the mAb epitope are most important. Such generalized methods for affinity chromatographic purification of antibodies rely on their interactions with ligands that bind non-variable regions on the antibody surface.[152,154] Usually these ligands are naturally occurring proteins that have this characteristic, like protein A from *Staphylococcus aureus* and protein G from *Streptococcus*.[155,156]

Affinity columns based on protein-A are the most widely used for several reasons. First, protein A is well-characterized and may be obtained from recombinant bacteria.[155,157] Second, protein A interacts with the Fc fragment of IgG with a high affinity constant (~10^{-7} M).[158] Third, protein A is stable over a wide pH range (2-11), it can refold after treatment with denaturating solutions like urea and guanidinium salts,[159] (it can even be "cleaned' using 0.5 M NaOH) hence the affinity supports can sometimes be recycled.[160] Finally, protein A is easily coupled to supports. Besides protein A, protein G is also a well-known natural binder of IgG, which has a broader binding range and higher affinity of IgG.[161] It binds all human IgG subclasses, and also mouse, rat and goat IgG.[162] The Fc fragment of IgG appears mainly responsible for the interaction with protein G, although a low degree interaction was also found for Fab fragments.[163] Protein G binds human IgG with a slightly higher dissociation constant (~10^{-8} M) than protein A (~10^{-7} M) does.[158]

Despite the research on affinity purification of IgGs, several important applications of this material are still restricted by the costs of purification. For the most common method, affinity chromatography using immobilized protein A or G, this relates to the cost of the native protein ligand. Moreover, leakage of protein A or G from the supports further complicates the process if very high states of purity are required.[164,165] For these reasons, inexpensive, synthetic ligands with appropriate affinities for the non-variable regions of IgG are interesting as potential ligands for affinity supports.

Prior work in this areas includes multimeric peptides,[153,166-169] triazine derivatives,[170-175] and some other molecules.[176] Some of these synthetic ligands showed good affinity with IgG and could be used to purify the target antibody with decent capacity and yield in high purity in a test system.[169,177,178] For example, a synthetic multibranched peptide ligand TG 19318 as a protein A mimetic had an apparent affinity constant with rabbit IgG of 3×10^5 M^{-1};[179] another important synthetic small molecule ligand 22/8**,** which is a triazine derivative designed and made by Lowe's group, mimicked protein A and was reported to have a K_a ~10^5 M^{-1} with IgG (Figure 5.2).[171] However, none of these compounds have found widespread commercial application. This may be due to various issues such as selectivity, capacity, and inappropriate binding constants.



ligand **22/8**

Figure 5.2. Structure of 22/8 and docking of 22/8 on IgG/protein A binding pocket.

Mabsorbent® A2P is a synthetic protein A mimetic that was commercialized by ProMetic Biosciences (UK).[180] It is composed of a di-substituted phenolic derivative of tri-chlorotriazine and is commercially available coupled to a 6% cross-linked agarose base matrix. The ligand was discovered from screening of ProMetic's combinatorial ligand libraries in binding IgG and was thought to mimic the structure of two key amino acid side-chains of protein A - Phe132 and Tyr133 that are found to play an important role in formation of the binary complex between protein A and the Fc fragment of IgG. It

binds all subclasses of IgG including IgG3, and could be used for the purification of polyclonal derived antibody therapeutics.[181]

5.2 Design and Synthesis of Cyclic Peptidomimetics of Protein A

5.2.1 Rationale

Crystallographic data for the complex between the key helix-loop-helix region of protein A (the "B-domain") and the human IgG Fc fragment (protein databank numbers:1l6x and 1FC2; Figure 5.3a) is thought to illustrate the key hot-spots for the interaction.[157] It features contact of residues in the two helices of protein A/fragment B (residues Gln128-Leu136, Glu144-Asp155) with β-turns of the IgG Fc region. This structural model is also supported by data from NMR[182,183] and mutational studies.[184] Consequently, it has been proposed[185] that the following protein A residues are important for binding the Fc region of IgG: Phe124, Phe132, Tyr133, Leu136, Ile150, Lys154, Gln128, Asn130, Asn147, and Arg146. Of these, Phe132 and Tyr133 may be pivotal; they form a hydrophobic pocket that encapsulates the Ile253 of IgG (Figure 5.3c). Moreover, this dipeptide structure motif has been found in four highly conserved regions of protein A, each of which is capable of binding the Fc fragment IgG from various species.[177]

a.

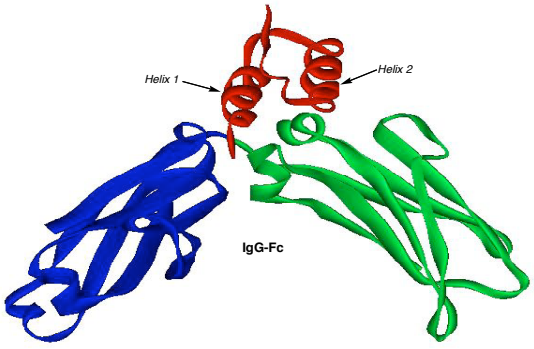
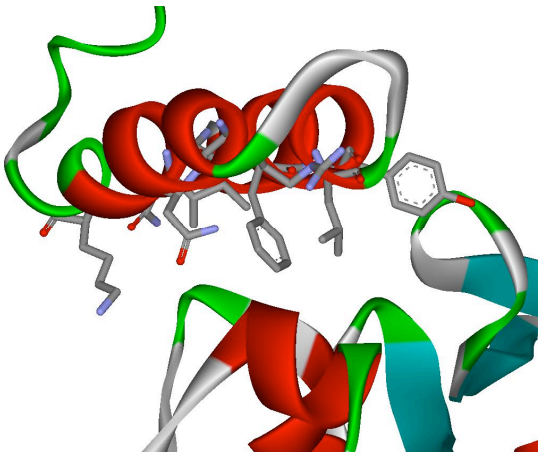


Figure 5.3. **a.** Fragment B of protein A complexed with IgG;[157] **b.** an expanded view of the same interaction from a different perspective; and, **c** the protein A helix-loop-helix region with the residues that make contact with protein A highlighted.

b.



c.

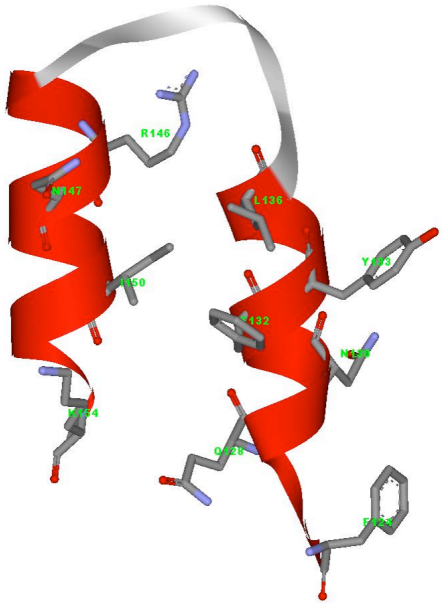
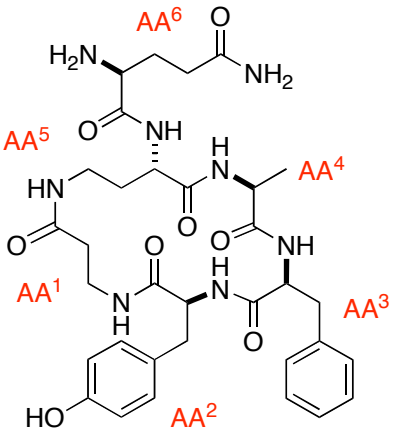


Figure 5.3. continued.

This research was performed to design peptidomimetics of the protein A helix-loop-helix region that would adopt similar conformations in solution, and which also retain the residues identified as hot-spots for binding to IgG. Consequently,

peptidomimetics **19a-i** were conceived (Figure 5.4). Tyrosine and phenylalanine motif have proved crucial in the binding studies; these residues were retained in each synthetic ligand. A Glutamine tail was used to help maintaining the desired conformation of the cyclic peptidomimetic. To examine the design, low energy conformers of compound **19a** were generated by quenched molecular dynamics (QMD) simulations[186,187] and compared with the complex of protein A and IgG (Figure 5.5a and b), also compound **19c** was docked with protein A fragment to prove their similarity (Figure 5.5c).

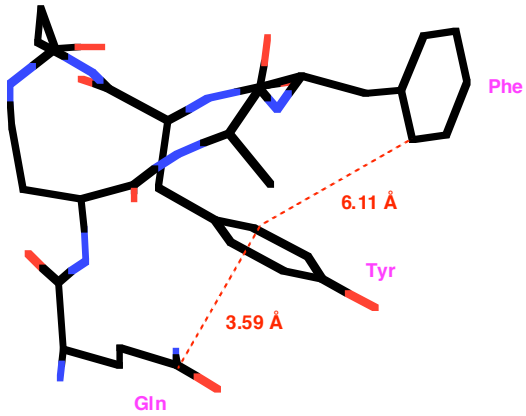


19

AA^1 = β-Ala, Ala, Gly
AA^2 = Phe, Tyr
AA^3 = Phe, Tyr
AA^4 = Ala
AA^5 = Dbu, Orn
AA^6 = Gln

Figure 5.4. Cyclic peptidomimetics **19a-i** of protein A.

a.



b.

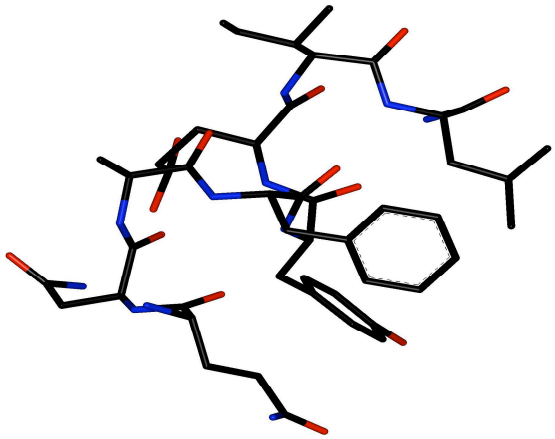


Figure 5.5. **a.** Low energy conformer of compound **19a** from quenched molecular dynamics (QMD) simulations; **b.** protein A fragment in complex with IgG; **c.** the best docking conformation of compound **19c** on protein A backbone.

c.

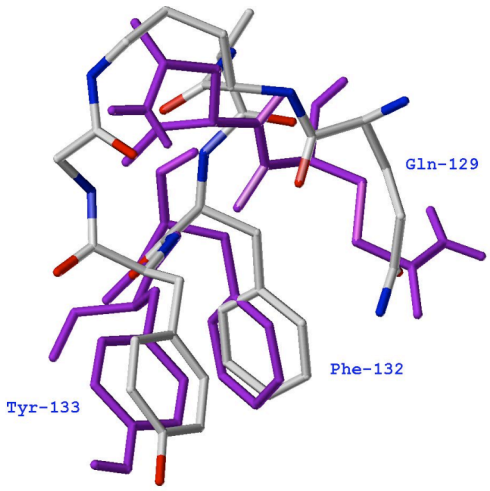


Figure 5.5. continued.

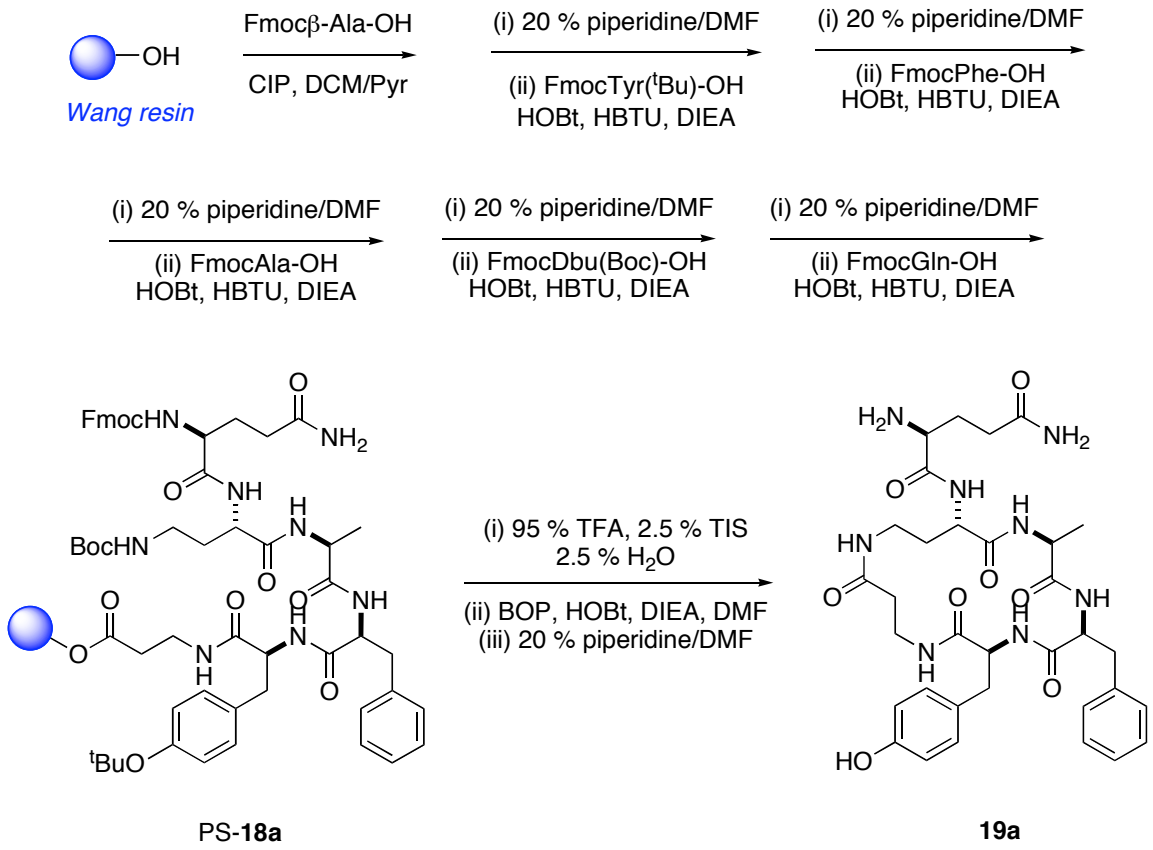
The backbone of each peptidomimetic **19** was constructed from a six-amino acid sequence in a cyclic format. Some amino acid residues were chosen for no particular reason except to construct the cyclic peptide scaffold. However, every compound features a combination of the two key amino acid residues for binding between protein A and IgG, Phe132-Tyr133. The combination can also be in reversed order, Tyr-Phe, or a close analog Phe-Phe. Glutamine was also involved in every compound for its potential function of maintaining a desired conformation.

5.2.2 Syntheses of Cyclic Peptidomimetics 19 of Protein A

The synthesis of cyclic peptidomimetic **19a** is outlined in Scheme 5.1 and used as an example. All the linear peptides were prepared via solid phase peptide synthesis (SPPS). Six Fmoc protected amino acids were coupled sequentially onto Wang resin following standard Fmoc peptide synthesis strategy. The linear hexapeptides **18a-i** were cleaved from the resin using TFA and the purity was evaluated by analytical HPLC, monitored by UV and SEDEX detector. All the linear peptides were found to have a satisfactory crude purity and were used in the next cyclization step directly.

A clean and efficient way of to perform the intramolecular cyclization of the linear peptide **18a-i** was not very easy to find because the long peptide chain made the intramolecular cyclization more difficult, and also the reaction solution needed to be dilute to avoid the formation of dimers, which required longer reaction times. Several conditions[188-190] were tested but gave unsatisfactory results because of epimerization formation of many byproducts, or slow and incomplete reaction. The best result was obtained when BOP reagent[191] was used along with HOBt, and DIEA as base in DMF at a concentration about 0.7 mM. After two days, the reaction was complete with a 100 % conversion from analytical HPLC. The cyclic compounds were purified via RP-preparative HPLC and the pure cyclic peptides were then subjected to 20 % piperidine in DMF to remove the Fmoc protection on glutamine to provide the final products **19a-i**, which were also purified via RP-preparative HPLC.

Scheme 5.1. Synthesis of cyclic hexapeptide **19a**.



PS-**18a** **19a**

A series of nine cyclic peptidomimetics of protein A **19a-i** was prepared; they were characterized by 1D and 2D NMR spectroscopes and ESI or MALDI-MS (Table 5.1).

Table 5.1. Summary of cyclic peptidomimetics **19a-i**.

comp'd **19**	amino acid sequences	UV purity[a] (%)	SEDEX purity[a] (%)	desired MS $(M+H)^+$	actual MS[b] $(M+H)^+$
a	Gln-Dbu-Ala-Phe-Tyr-β-Ala	100	100	681.33	681.33
b	Gln-Dbu-Ala-Phe-Tyr-Ala	100	100	681.33	681.32
c	Gln-Orn-Ala-Phe-Tyr-Gly	100	100	681.33	681.35
d	Gln-Dbu-Ala-Tyr-Phe-β-Ala	100	100	681.33	681.27
e	Gln-Dbu-Ala-Tyr-Phe-Ala	100	100	681.33	681.21
f	Gln-Orn-Ala-Tyr-Phe-Gly	100	100	681.33	681.26
g	Gln-Dbu-Ala-Phe-Phe-β-Ala	100	100	665.33	665.30
h	Gln-Dbu-Ala-Phe-Phe-Ala	100	100	665.33	665.33
i	Gln-Orn-Ala-Phe-Phe-Gly	100	100	665.33	665.32

[a] determined after HPLC purification.
[b] Mass of **19a-c** and **h** was obtained from ESI ionization method, the rest was from MALDI.

These compounds were sent to Amersham in Sweden for biological tests. However, no good binding data with IgG was obtained from these molecules.

5.3 Design and Synthesis of Small Molecule Mimetics of Protein G

5.3.1 Rationale

The focus of this research was to find small ligands based on hot-spots in protein G that are known as key sites for interaction with IgG. The essential features of the protein G:IgG interaction are as follows.

Protein G has several structural segments, but the one called "domain III" is thought to contain the key hot spots for binding IgG.[192-194] This domain interacts with both the Fab and Fc regions of IgG. Crystal structures of the protein G domain III complexed with the Fab (PDB: 1IGC)[193] and with the Fc fragments (PDB: 1FCC)[195] of IgG are known. Figure 5.6 shows details of the protein G domain III:Fab interaction. This features a sheet region and a helical region on the protein G fragment in close contact with the C_H1 domain of Fab. The first of these interactions is an antiparallel alignment of the second β-strand of the protein G domain III with the seventh β-strand of C_H1 domain of Fab (Figure 5.6b). The residues involved in this contact region are Lys15A to Thr22A from protein G, and Ser209H to Lys216H from the IgG Fab fragment. This β/β-sheet interaction site incorporates the charged residues, Lys18A, Glu20A on protein G interaction with Lys212H and Asp214H on IgG, which lie along the outer surface of the β-strand, exposed to the solvent environment. The other key interaction is between Tyr38A to Gly43A in the *C*-terminal α-helix of protein G domain III, and Pro125H to Tyr129H on the first β-strand of the IgG C_H1 region (Figure 5.6c). We hypothesized that these two interactions form a hydrophilic perimeter around a hydrophobic core formed among Pro125H, Thr211H and Val213H. This is a typical type of hot-spot in protein-protein interactions that has been compared with an "O-ring".[196] We chose to target this region of the IgG Fab fragment for our peptidomimetic design. (Some target molecules with similar structures have been docked in the region of IgG that binds the protein G domain III using computer program AFFINITY to show the potential binding modes; these are not shown.)

Dendric molecules like compound **DAB-Am-4** are attractive frameworks for syntheses of tetravalent and higher oligomers of biologically active ligands. This is important because many protein-protein interactions feature multivalent ligands

interacting with receptors to initiate or perturb their association.[46,197-201] Within this area of interest there are many situations in which it would be advantageous to attach ligands to all the sites except one on a dendric system; for example, if three out of the four primary amine groups in **DAB-Am-4** were coupled to biomolecules then the remaining unfunctionalized one could support a fluorescent or similar label, or be used to attach it to an affinity support. Further, in combinatorial chemistry, size exclusion methods may facilitate more facile isolation of dendric molecules than similar monofunctionalized ones.[202,203] For this particular case, it was anticipated that one unit on a dendrimer might rest comfortably in the hydophobic pocket while the others could form additional contacts with the protein surface. The site on the dendric molecule that did not contain a peptiomimetic "warhead" would be used to anchor the whole dendric molecule to a solid support.

a.

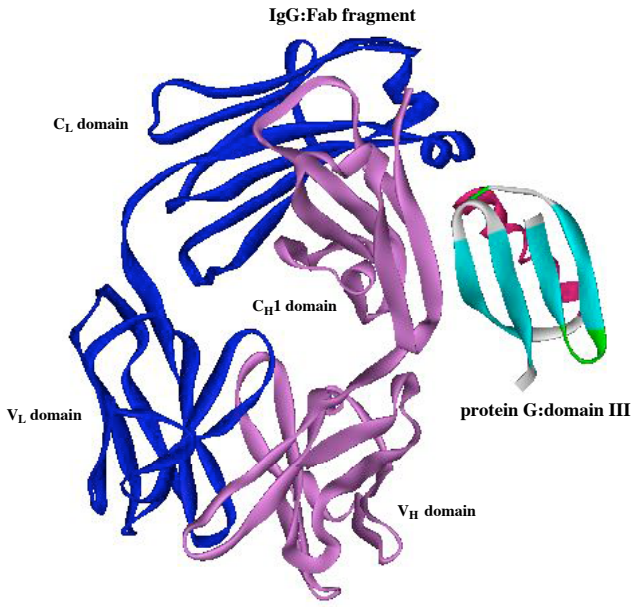
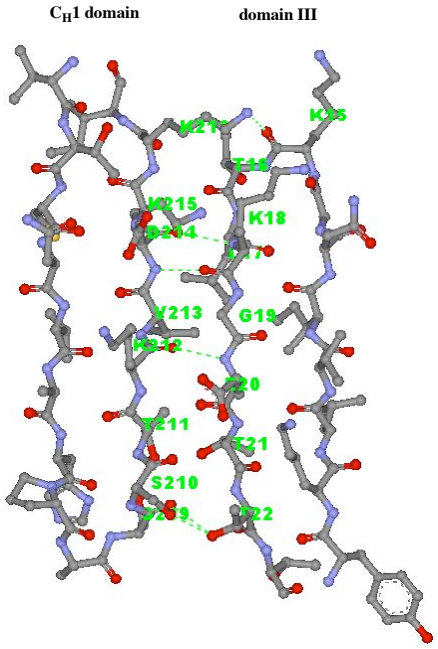


Figure 5.6. ViewerPro plot of interaction details between domain III of protein G and the C_H1 domain of Fab. Hydrogen bonds between these two domains are marked using dashed lines. **a.** the overall interaction; **b.** the contact between the Fab and protein G β-sheets; and **c.** is the interaction between a β-strand of Fab (on the left) and the α-helix of domain III (on the right).

b.



c.

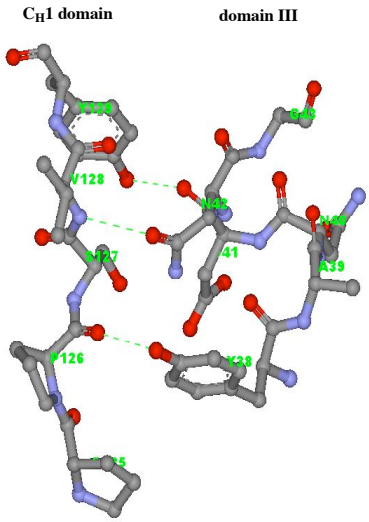


Figure 5.6. continued.

5.3.2 Synthesis of Small Molecule Mimetics 22 and 23 of Protein G

A set of 14 small molecule ligands **22a-k** and **23a-c** (Figure 5.7) was prepared in solution phase. The concept was that after biology test, the active monomers will be selected from the library and used as monomeric ligands on **DAB-Am-4** polyamine to make the trivalent molecules **24a-d** with a free handle for further biological assays applying the prior methodology developed in our group (not published).

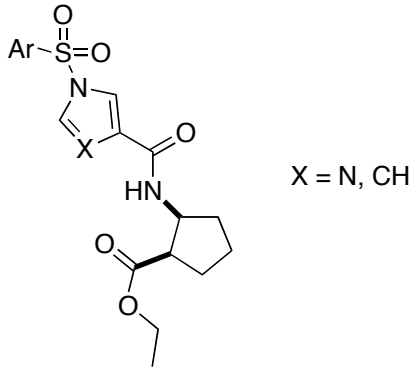


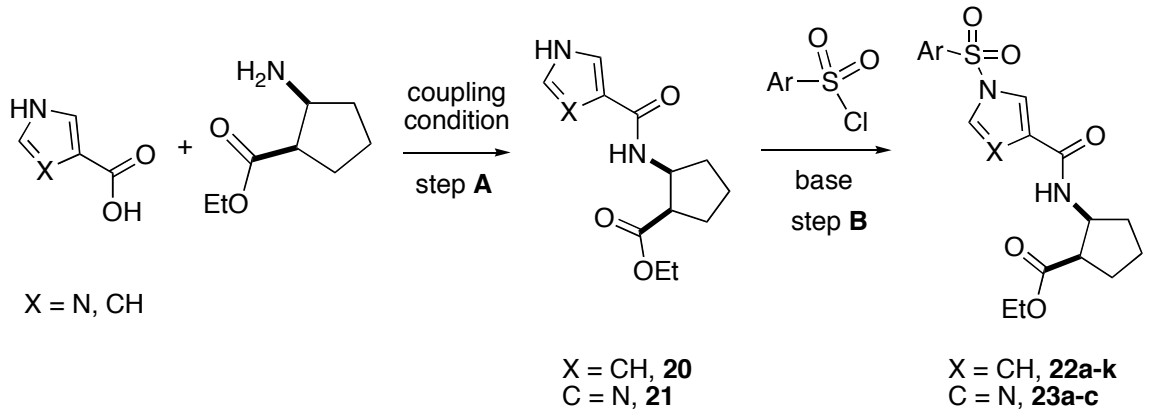
Figure 5.7. Heterocyclic small molecules **22** and **23** as protein G mimetics.

The target molecules consisted of three components, aryl sulfonyl chloride, pyrrole or imidazole carboxylic acid and a *cis*-amino ester. The syntheses stared from the only conserved fragment, the amino ester, which was coupled with the heterocyclic carboxylic acid to form the amide **20** and **21**. Then different aryl sulfonyl chlorides were reacted with the coupling adducts under basic condition to afford each target compound **22** and **23**(Scheme 5.2).

The coupling conditions for the pyrrole compounds and imidazole compounds were slightly different based on their properties. We used DIC/DMAP for the pyrrole coupling and the desired product **20** was recrystallized from ethanol. Coupling condition using EDC/HOBt/NMM was found to work better for the imidazole acid and the crude product **21** in high purity was used in the next step directly. Different bases were used for deprotonation because of the difference in pKa of the pyrrole N<u>H</u> and imidazole N<u>H</u>. A stronger base (*i.e.* NaOH)[204] was used for pyrrole and Et_3N[205] worked well for

imidazole in the nucleophilic substitution reaction with aryl sulfonyl chloride to yield **22** and **23**, respectively. All the final products were purified via flash chromatography.

Scheme 5.2. Synthetic scheme of pyrrole and imidazole derivatives.



conditions:

X = CH, step **A** = DIC, DMAP, CH_2Cl_2/DMF, 0 °C
step **B** = NaOH, dicholoroethane, 0 °C to 25 °C

X = N, step **A** = EDCI, HOBt, NMM, DMF, 25 °C
step **B** = Et_3N, THF, 25 °C

We made a total of eleven pyrrole derivatized compounds **22a-k** and three imidazole derivatives **23a-c** following the above scheme. Their structures and the yields were summarized in Table 5.2.

Table 5.2. Summary of pyrrole and imidazole derivatized ligands **22** and **23**.

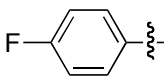
comp'd **22** and **23**	Ar	X	yield[a] (%)
22a		CH	43

Table 5.2. continued.

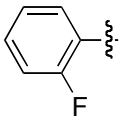
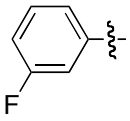
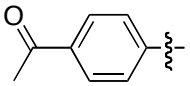
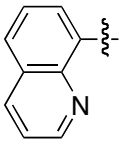
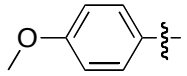
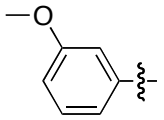
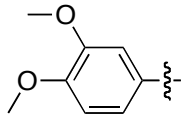
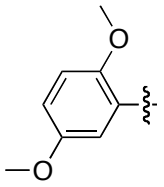
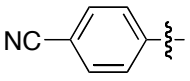
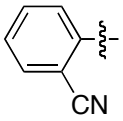
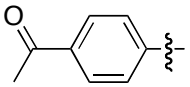
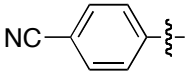
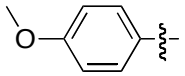
comp'd **22** and **23**	Ar	X	yield[a] (%)
22b		CH	69
22c		CH	45
22d		CH	44
22e		CH	71
22f		CH	90
22g		CH	45
22h		CH	73
22i		CH	84
22j		CH	44
22k		CH	33
23a		N	67
23b		N	65

Table 5.2. continued**.**

comp'd **22** and **23**	Ar	X	yield[a] (%)
23c		N	80

[a] Yields of pyrrole derivatives **22a-k** are calculated based on the second step **B**; the yields of imidazole derivatives **23a-c** are calculated based on two steps **A** and **B**.

5.3.3 SAR Screening of Small Molecular Protein G Mimetics 22 and 23

All 14 compounds **22a-k** and **23a-c** were sent to Sweden along with other protein G mimetics made by another group member for SAR screening.

All ligands were dissolved in 100% DMSO to a concentration of 50 mM. The stock solutions were then further diluted to the concentration of 250 uM. The dilution was done in such a way that the concentration of DMSO in the samples was carefully matched to 5%. Eleven of the compounds had a low solubility at this concentration, and one compound **22h** was totally precipitated.

A CM5 sensor chip was used and a monoclonal Fab κ fragment was immobilized to a level of 4000 RU in flow cell 2 (Fc2). Fc1 was left totally unmodified and used in the DMSO calibration during the evaluation. PBS with 5% of DMSO was used as running buffer during the assay.

The compounds were injected one by one and in order to reduce carry over effects between sample injections. The experiment was carefully designed by including wash steps and buffer injections in addition to the sample injections. The protein surface was also regenerated with an injection of 1 M ethanolamine, pH 8.5 after each sample injection. Every sample was injected in duplicates over the two surfaces (Fc1 and Fc2).

The soluble compounds were first tested in duplicates and then the compounds that showed less solubility were tested. Buffer injections were also made during the assay and warfarin was used as a negative control ligand. Protein G, used as a positive control, was injected every 10^{th} cycle in order to measure that the protein surface

maintained activity. It was observed that the protein G responses showed only a slight decrease during the assay, which indicated that the protein surface remained active. The signals from these sample injections were shown in Table 5.3.

Table 5.3. SAR signals from sample injections.

comp'd **22** and **23**	M. W.	normalised response: replicate 1	normalised response: replicate 2	D1: replicate 1	D2: replicate 1	D1: replicate 2	D2: replicate 2
22c	408.12	4.58	1.76	1.4	1.2	0.4	0.4
22d	432.14	14.75	14.14	5.1	4.7	5	4.8
22i	450.15	11.13	8.60	3.6	3.4	2.8	2.7
22k	415.12	9.78	9.16	3.1	2.8	3.2	3.1
22j	415.12	9.30	12.54	2.8	2.5	2.4	2.3
23b	416.12	2.56	4.22	-0.5	-0.5	-0.3	-0.4
23c	421.13	0.86	2.82	0.3	-0.1	0	0
23a	433.13	2.59	3.89	1	0.9	0.8	0.4
22a	408.12	4.63	4.03	0.6	0.3	1.7	1.6
22b	408.12	-0.11	0.46	-0.3	-0.5	0.6	0.8
22e	441.14	3.87	-1.24	0.5	0.3	0	-0.2
22f	420.14	5.00	-0.62	-0.2	-0.5	0.1	0.4
22g	420.14	2.00	-0.14	-0.3	-0.3	0.5	0.5

Table 5.3 shows the reference subtracted (Fc2-1) and calibrated sample responses. The responses have also been normalized according to molecular weight, enabling the comparison of ligands of different sizes. In general, the sample injections generated low SPR responses in the range 0-6.4 RU (calibrated, not normalized). The compounds that generated the highest normalized signals, which indicating highest affinity towards a Fab

fragment were **22d**, **22i**, **22j** and **22k** (marked yellow in Table 5.3). The response in Fc2-1 after the sample injection was also registered in two dissociation points. D1 is registered 10 seconds after end of sample injection and D2 at 20 seconds after termination of injection. Ligands **22d**, **22i**, **22j** and **22k** (marked yellow) also showed high dissociation signals (not normalized), which indicate a slow dissociation. Unfortunately, the data was only collected from injections of a single concentration; therefore no affinity constants could be calculated.

5.3.4 Synthesis of Multivalent Mimetics 25 from Active Monovalent Ligands 22

To increase the binding affinity, we attached the 4 active ligands **22d**, **22i**, **22j** and **22k** selected from SAR screening to a dendrimer core, **DAB-Am-4** to generate 4 corresponding multivalent mimetics **25a-d** (Figure 5.8).

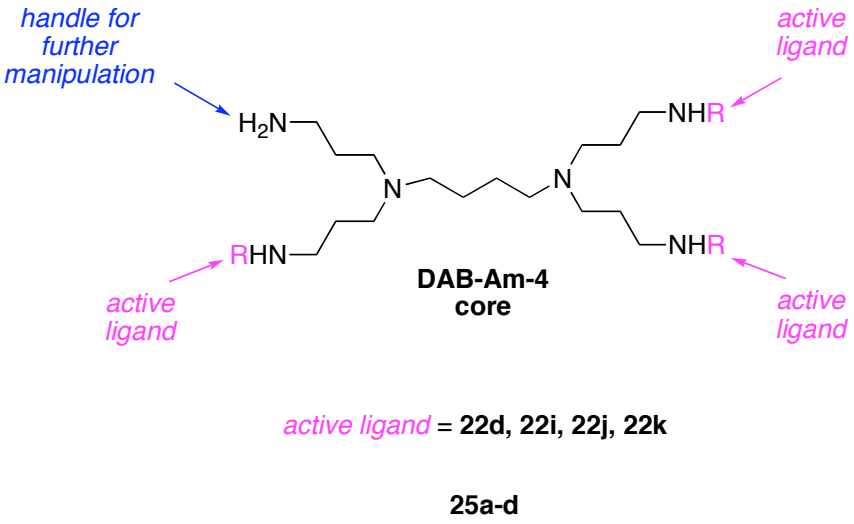


Figure 5.8. Structure representation of active ligand fuctionalized dendrimers **25a-d** on a **DAB-Am-4** core.

The whole process is outlined in Scheme 5.3. First, each monomer ester was hydrolyzed to a carboxylic acid. To avoid the possible retro-reaction of forming the sulfonamide bond under basic condition, the esters were hydrolyzed under mild acidic conditions. At 40 °C the hydrolization proceeded smoothly without racemization and

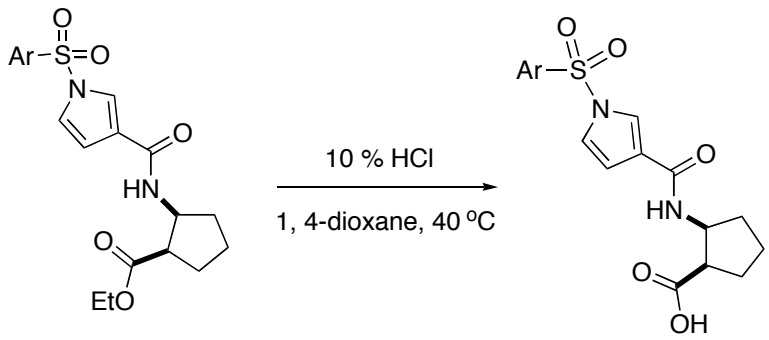
offered the desired carboxylic acid **24a-d** in high yield after flash chromatography (Scheme 5.3a).

To selectively use $n - 1$ functional groups out of a total of n at the periphery of a dendric entity, we used the methodology developed in our group as depicted in Scheme 5.3b, which takes advantage of the relatively disperse arrangement of functional groups on supports for solid-phase syntheses.[206]

TentaGel S PHB resin (which is functionalized with the Wang acid linker) was further derivatized with 4-nitrophenyl chlorocarbonate to give the electrophilic resin.[207] This was reacted with **DAB-Am-4**, coupled with the monomeric carboxylic acid ligand **24a-d**. In any event, after considerable experimentation, we found that the efficiency of the synthesis with respect to formation of the desired product was highly dependent on the concentration of the **DAB-Am-4** used in the ligand coupling reaction. To satisfy both the coupling efficiency and ease of handling (*eg* washing), the optimal concentration we used was 1.5 M; if higher concentrations (> 2 M) were used then the reaction liquid phase became quite viscous and the coupling efficiency declined. Finally, the trivalent compounds were cleaved from the resin using TFA (Scheme 5.3b).

Scheme 5.3. Synthesis of trivalent ligands **25a-d. a.** hydrolysis of ester **22d**, **22i**, **22j** and **22k**; **b.** attachement of monovalent ligand **24a-d** to a **DAB-Am-4** core.

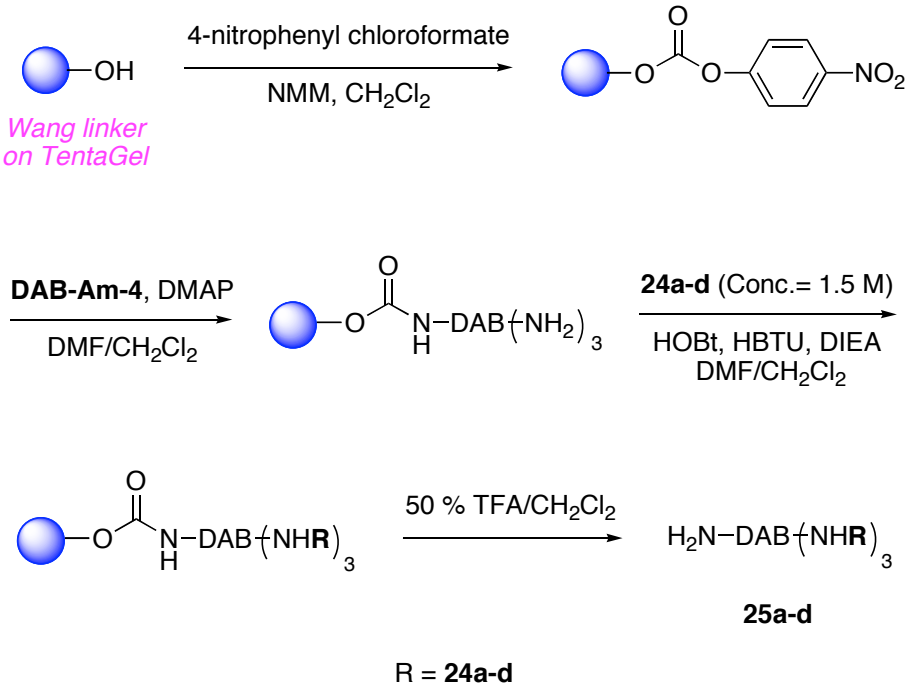
a.



22d, 22i, 22j, 22k **24a-d**

Scheme 5.3. continued.

b.



The final products **25a-d** were purified via RP-preparative HPLC, and identified via MALDI-MS. Table 5.4 outlines details of the purities of the material cleaved from the resin, and the yields of the desired products are based on the loading of the resin.

Table 5.4. Summary of the active ligand-functionalized dendrimers **25a-d.**

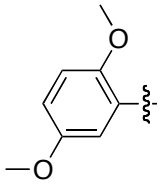
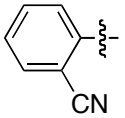
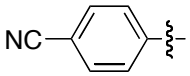
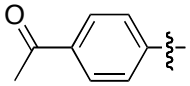
trivalent compound **25**	Monovalent ligand functionality **R**	Ar (in **R**)	X (in **R**)	yield (%)
a	**i**		CH	20

Table 5.4. continued.

trivalent compound **25**	monovalent ligand functionality **R**	Ar (in **R**)	X (in **R**)	yield (%)
b	**k**		CH	18
c	**j**		CH	21
d	**d**		CH	24

Since the attempts to develop a methodology of coupling synthetic ligands to a higher homolog of **DAB-Am-4** having more amine groups were unencouraging, we only made these four functionalized dendrimers. Unfortunately, these compounds were never sent to other places for testing, so their activities and affinities still remain unrevealed.

5.4 Summary

IgG, as the most commonly used antibody in immunochemical procedures, has widespread applications in various area such as diagnostics,[139,140] pharmaceutical purifications, and perhaps the most exacting one, therapeutics, [140,142-145,146,147] in which the purity of IgGs is particularly critical. Therefore, development of time- and cost-effective techniques for large-scale isolation and purification of IgG are of great importance.

One of the most widely applied purification technique of IgG is affinity chromatography.[141,151,152] Monoclonal antibodies (mAbs) designed to bind a specific antigen may be purified using supports based on that substance.[153] Usually the affinity ligands used for this purpose are naturally occurring proteins that bind non-variable regions on the antibody surface,[152,154] like *Staphylococcus aureus* protein A and *Streptococcus* protein G,[155,156] which both have a high affinity constant ($\geq 10^{-7}$M) with

IgG (protein A binds the Fc fragment and protein G binds the Fc and Fab fragment of IgG) [2,158]. However, due to the cost and other practical reasons, inexpensive, appropriate synthetic ligands of IgG are extremely interesting as potential ligands for affinity supports. Prior work in this area did not provide many compounds of widespread commercial application.

In a collaborative project with Amersham company, we prepared a series of mimetics of protein A and protein G from careful and reasonable design based on their interactions with IgG. The targets were usually generated from molecular modeling studies of the protein-ligand interaction information obtained from their crystal complex.

A set of nine cyclic hexapeptides **19a-i** was prepared from SPPS and solution phase intramolecular macrocyclization reaction. All the peptidomimetics feature the dipeptide motif, "Phe132-Tyr133" (or Tyr-Phe, Phe-Phe), which was found pivotal in the protein A and IgG binding event. However, no interesting binding result was obtained for these compounds.

For protein G, we prepared a total of 14 pyrrole or imidazole derivatized small molecule ligands **22a-k** and **23a-c** with aromatic functionalities through efficient and straightforward syntheses. These synthetic ligands were tested in SAR screening with IgG Fab fragment and 4 of them (**22d**, **22i**, **22j** and **22k)** showed relatively higher affinity. Therefore we attached these selected ligands to a dendrimer core, **DAB-Am-4** and generated 4 trivalent mimetics **25a-d** (with a free handle); we were hoping this would increase the binding affinity. Unfortunately, these compounds were never tested.

Overall, we made some novel synthetic ligands as natural protein A and protein G mimetics for affinity chromatography of IgG purification. They were tested in binding assays and several showed some kind of binding affinity with Fab fragment of IgG in SAR screening, which may be useful for further development in this area.

CHAPTER VI

CONCLUSIONS

Peptidomimetics are useful tools in studying protein-protein interaction events via mimicry of the crucial interaction sites. They are typically used to mimic or disrupt the structural environment at hot-spots. Heterocyclic building blocks can be useful as scaffolds for peptide backbone modifications. They import rigidity and enable the functionalized side-chain substituents to adopt relatively constrained conformations. Heterocyclic dipeptide mimetics can be accessed from conventional procedures and show various functionalities. They are used as protein secondary structure mimics and for structure modification. In chapter I, strategies in design and synthesis of heterocyclic amide bond surrogates of dipeptide mimetics was discussed.

In general, we think the incorporation of the amino acid side-chains, particularly those that occur most frequently at hot-spots, in the mimics is quite important; and good mimics should also be somewhat flexible to be able to access favorable conformations that are close to those of the parental peptides in binding protein targets.

Our laboratory has been involved in the design and synthesis of small molecules to mimic surface exposed β-turn regions of NGF (and other neurotrophins). Specifically, we have targeted mimics of the NGF turn regions that seem to be involved in those "hot-spot" interactions between NGF and TrkA.

Given the interesting biochemical and biological profile of previously reported peptidomimetic **D3** and the notion that changing a small molecule from a monovalent to bivalent form could generate more potent agonists for RTK ligands, we prepared a set of bivalent molecules **1** using a novel combinatorial approach. A close analog of **D3** (*i.e.* **A**) was synthesized on solid support and attached to different length linkers. They were then paired on a triazine scaffold via our previously published methodology, to give a small library of bivalent compounds representing three linker fragments in all permutations.

One homo-bivalent compound, **1-ss** showed higher affinity than others with TrkA and a ~50-fold better affinity than the parental monovalent peptidomimetic (~200 nM

versus ~10 μM), and retains high selectivity towards TrkA in binding assays. It can block the binding of another more efficient ligand, anti-TrkA mAb 5C3, and also block NGF binding. In biological assays, **1-ss** was found to antagonize ligand-dependent activation of TrkA as a pure receptor antagonist while the parental monovalent peptidomimetic is an agonist towards NGF, which is intriguing. This is probably because the bivalent mimetic binds to two receptors, and prevents conformational changes of the receptor dimer, which need to be induced by an agonistic ligand. Unfortunately, conformational changes in TrkA upon the binding of its protein ligand NGF have not yet been demonstrated through structural analyses, and this is still an issue in the current debate.

The approach to a set of monovalent diketopiperazine mimetics **4-7** was illustrated. Intramolecular S_N2 cyclization of corresponding dipeptides in solution yielded the desired mimetics efficiently. These DKP compounds contain two side-chain functionalities that are both derived from natural amino acids, which were chosen based on the $i + 1$ and $i + 2$ residues of the turn regions of NGF. Each monomer was attached to a short or long linker before dimerization. A library of 45 biotin-labeled bivalent molecules **9** was obtained via the efficient combinatorial strategy described above with some modification.

In primary screening, four bivalent compounds exhibited relatively higher binding affinity and selectivity against TrkC-expressing cells at 50 uM in FACScan assay, and they all contain at least one cyclo(Ile-Lys)-long linker fragment. In cell survival assay, three bivalent compounds blocked the trophic activity of NT-3 in TrkC cells at 10 uM, and no significant effect was observed for them in IGF-1R and TrkA cells. These results need to be further confirmed by repeating with doses. Interestingly, compound **9gg,** which also showed preferential binding to TrkC receptors, is a homo-bivalent molecule of two cyclo(Ile-Lys)-long linker moieties.

Parallel solution phase synthesis allows quick and efficient generation of a set of 1,3,4-oxadizole based dipeptide mimetics **12** in good yields on gram scale. These heterocyclic compounds feature various natural amino acid side-chain functionalities. In attempts to assemble the monovalent mimics into bivalent molecules **15** via the above dimerization method by simply manipulating the solvent systems, we found that some

reactions did not proceed cleanly, which perhaps is due to some certain reactive groups in the molecule based on our test results. The problems were partially solved but more studies need to be done. The biological activities of the oxadiazole molecules **12** and **17** are under investigation. So far, NIH screening centers have reported six compounds showing certain activities in four bioassays of different types.

In a collaborative project with Amersham company, two different peptidomimetic types that resemble protein A and protein G binding regions were generated and tested as binding factors in affinity columns for purification of IgG. Cyclic hexapeptides **19** as protein A mimics feature the dipeptide motif, "Phe132-Tyr133", which was found pivotal in the binding event. However, no interesting binding result was obtained for these compounds. Heterocycle-based small molecules **22** and **23** featuring a variety of aromatic functionalities were synthesized as protein G mimics. Four compounds in this library showed some affinity towards a Fab fragment of IgG in SAR screening, thus they were attached to a dendrimer core to give four multivalent mimetics **25** to obtain higher binding affinity. These compounds have not been tested yet.

REFERENCES

1. von Roedern, E. G.; Lohof, E.; Hessler, G.; Hoffmann, M.; Kessler, H., *J. Am. Chem. Soc.* **1996**, *118*, 10156-10167.

2. Gruner, S. A. W.; Locardi, E.; Lohof, E.; Kessler, H., *Chem. Rev.* **2002**, *102*, 491-514.

3. Risseeuw, M. D. P.; Overhand, M.; Fleet, G. W. J.; Simone, M. I., *Tetrahedron: Asymmetry* **2007**, *18*, 2001-2010.

4. Chakraborty, T. K.; Ghosh, S.; Jayaprakash, S.; Sharma, J. A. R. P.; Ravikanth, V.; Diwan, P. V.; Nagaraj, R.; Kunwar, A. C., *J. Org. Chem.* **2000**, *65*, 6441-6457.

5. Baldwin, J. E.; Hulme, C.; Schofield, C. J.; Edwards, A. J., *Journal of the Chemical Society, Chemical Communications* **1993**, *11*, 935-936.

6. Vojkovsky, T.; Weichsel, A.; Patek, M., *J. Org. Chem.* **1998**, *63*, 3162-3163.

7. Lewis, J. G.; Bartlett, P. A., *J. Comb. Chem.* **2003**, *5*, 278-284.

8. Polyak, F.; Lubell, W. D., *J. Org. Chem.* **2001**, *66*, 1171-1180.

9. Qiu, W.; Gu, X.; Soloshonok, V. A.; Carducci, M. D.; Hruby, V. J., *Tetrahedron Lett.* **2001**, *42*, 145-148.

10. Ndungu, J. M.; Gu, X.; Gross, D. E.; Cain, J. P.; Carducci, M. D.; Hruby, V. J., *Tetrahedron Lett.* **2004**, *45*, 4139-4142.

11. Ndungu, J. M.; Gu, X.; Gross, D. E.; Ying, J.; Hruby, V. J., *Tetrahedron Lett.* **2004**, *45*, 3245-3247.

12. Ndungu, J. M.; Cain, J. P.; Davis, P.; Ma, S.-W.; Vanderah, T. W.; Lai, J.; Porreca, F.; Hruby, V. J., *Tetrahedron Lett.* **2006**, *47*, 2233-2236.

13. Borg, S.; Vollinga, R. C.; Labarre, M.; Payza, K.; Terenius, L.; Luthman, K., *J. Med. Chem.* **1999**, *42*, 4331-4342.

14. Brik, A.; Yang, Y.-Y.; Ficht, S.; Wong, C.-H., *J. Am. Chem. Soc.* **2006**, *128*, 5626-5627.

15. Appendino, G.; Bacchiega, S.; Minassi, A.; Cascio, M. G.; De Petrocellis, L.; Di Marzo, V., *Angew. Chem., Int. Ed.* **2007**, *46*, 9312-9315.

16. Tam, A.; Arnold, U.; Soellner, M. B.; Raines, R. T., *J. Am. Chem. Soc.* **2007**, *129*, 12670-12671.

17. von Geldern, T. W.; Hutchins, C.; Kester, J. A.; Wu-Wong, J. R.; Chiou, W.; Dixon, D. B.; Opgenorth, T. J., *J. Med. Chem.* **1996**, *39*, 957-967.

18. von Geldern, T. W.; Kester, J. A.; Bal, R.; Wu-Wong, J. R.; Chiou, W.; Dixon, D. B.; Opgenorth, T. J., *J. Med. Chem.* **1996**, *39*, 968-981.

19. Mann, E.; Kessler, H., *Org. Lett.* **2003**, *5*, 4567-4570.

20. Biron, E.; Chatterjee, J.; Kessler, H., *Org. Lett.* **2006**, *8*, 2417-2420.

21. Du, W.; Hagmann, W. K.; Hale, J. J., *Tetrahedron Lett.* **2006**, *47*, 4271-4274.

22. Heller, S. T.; Natarajan, S. R., *Org. Lett.* **2006**, *8*, 2675-2678.

23. Holmes, C. P.; Chinn, J. P.; Look, G. C.; Gordon, E. M.; Gallop, M. A., *J. Org. Chem.* **1995**, *60*, 7328-7333.

24. Evindar, G.; Batey, R. A., *Org. Lett.* **2003**, *5*, 1201-1204.

25. Campiglia, P.; Gomez-Monterrey, I.; Carotenuto, A.; Lama, T.; Diurno, M. V.; Bertamino, A.; Mazzoni, O.; Sala, M.; Novellino, E.; Grieco, P., *Letters in Organic Chemistry* **2006**, *3*, 539-545.

26. Chaudhry, P.; Schoenen, F.; Neuenswander, B.; Lushington, G. H.; Aube, J., *J. Comb. Chem.* **2007**, *9*, 473-476.

27. Noth, J.; Frankowski Kevin, J.; Neuenswander, B.; Aube, J.; Reiser, O., *Journal of Combinatorial Chemistry* **2008**, *10*, 456-459.

28. Zeng, Y.; Li, Q.; Hanzlik, R. P.; Aube, J., *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3034-3038.

29. Goodfellow, V. S.; Laudeman, C. P.; Gerrity, J. I.; Burkard, M.; Strobel, E.; Zuzack, J. S.; McLeod, D. A., *Molecular Diversity* **1996**, *2*, 97-102.

30. Burkholder, T. P.; Huber, E. W.; Flynn, G. A., *Bioorg. Med. Chem. Lett.* **1993**, *3*, 231-234.

31. Lena, G.; Lallemand, E.; Gruner, A. C.; Boeglin, J.; Roussel, S.; Schaffner, A.-P.; Aubry, A.; Franetich, J.-F.; Mazier, D.; Landau, I.; Briand, J.-P.; Didierjean, C.; Renia, L.; Guichard, G., *Chem. Eur. J.* **2006**, *12*, 8498-8512.

32. Olson, G. L.; Voss, M. E.; Hill, D. E.; Kaln, M.; Madison, V. S.; Cook, C. M., *J. Am. Chem. Soc.* **1990**, *112*, 323-333.

33. Guzzo, P. R.; Trova, M. P.; Inghardt, T.; Linschoten, M., *Tetrahedron Lett.* **2001**, *43*, 41-43.

34. Sarabu, R.; Bolin, D. R.; Campbell, R.; Cooper, J. P.; Cox, D.; Gaizband, D.; Makofske, R.; Nagy, Z.; Olson, G. L., *Drug Design and Discovery* **2002**, *18*, 3-7.

35. Grieco, P.; Campiglia, P.; Gomez-Monterrey, I.; Novellino, E., *Tetrahedron Lett.* **2002**, *43*, 1197-1199.

36. Grieco, P.; Giusti, L.; Carotenuto, A.; Campiglia, P.; Calderone, V.; Lama, T.; Gomez-Monterrey, I.; Tartaro, G.; Mazzoni, M. R.; Novellino, E., *J. Med. Chem.* **2005**, *48*, 3153-3163.

37. Borg, S.; Estenne-Bouhtou, G.; Luthman, K.; Csoregh, I.; Hesselink, W.; Hacksell, U., *J. Org. Chem.* **1995**, *60*, 3112-3120.

38. Chagnault, V.; Lalot, J.; Murphy, P. V., *ChemMedChem* **2008**, *3*, 1071-1076.

39. Oh, K.; Guan, Z., *Chem. Commun.* **2006**, *29*, 3069-3071.

40. Crescenza, A.; Botta, M.; Corelli, F.; Santini, A.; Tafi, A., *J. Org. Chem.* **1999**, *64*, 3019-3025.

41. Shao, H.; Lee, C.-W.; Zhu, Q.; Gantzel, P.; Goodman, M., *Angew. Chem., Int. Ed.* **1996**, *35*, 90-92.

42. Tamura, S. Y.; Semple, J. E.; Reiner, J. E.; Goldman, E. A.; Brunck, T. K.; Lim-Wilby, M. S.; Carpenter, S. H.; Rote, W. E.; Oldeshulte, G. L.; Richard, B. M.; Nutt, R. F.; Ripka, W. C., *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1543-1548.

43. Angell, Y.; Chen, D.; Brahimi, F.; Saragovi, H. U.; Burgess, K., *J. Am. Chem. Soc.* **2008**, *130*, 556-565.

44. Kaplan, D. R.; Miller, F. D., *Current Opinion in Cell Biology* **1997**, *9*, 213-221.

45. Mischel, P. S.; Umbach, J. A.; Eskandari, S.; Smith, S. G.; Gundersen, C. B.; Zampighi, G. A., *Biophys. J.* **2002**, *83*, 968-976.

46. Heldin, C.-H., *Cell* **1995**, *80*, 213-223.

47. Schlessinger, J., *Cell* **2002**, *110*, 669-672.

48. Ullrich, A.; Schlessinger, J., *Cell* **1990**, *61*, 203-212.

49. Maliartchouk, S.; Saragovi, H. U., *J. Neurosci.* **1997**, *17*, 6031-6037.

50. Maliartchouk, S.; Feng, Y.; Ivanisevic, L.; Debeir, T.; Cuello, A. C.; Burgess, K.; Saragovi, H. U., *Molec. Pharmacol.* **2000**, *57*, 385-391.

51. McInnes, C.; Sykes, B. D., *Biopolymers FIELD Full Journal Title:Biopolymers* **1997**, *43*, 339-366.

52. Lin, B.; Li, Z.; Park, K.; Deng, L.; Pai, A.; Zhong, L.; Pirrung, M. C.; Webster, N. J. G., *Journal of Pharmacology and Experimental Therapeutics* **2007**, *323*, 579-585.

53. Schlein, M.; Ludvigsen, S.; Olsen, H. B.; Andersen, A. S.; Danielsen, G. M.; Kaarsholm, N. C., *Biochem.* **2001**, *40*, 13520-13528.

54. Wehrman, T.; He, X.; Raab, B.; Dukipatti, A.; Blau, H.; Garcia, K. C., *Neuron* **2007**, *53*, 25-38.

55. Feng, Y.; Pattarawarapan, M.; Wang, Z.; Burgess, K., *Org. Lett.* **1999**, *1*, 121-124.

56. Reyes, S.; Pattarawarapan, M.; Roy, S.; Burgess, K., *Tetrahedron* **2000**, *56*, 9809-9818.

57. Pattarawarapan, M.; Chen, J.; Steffensen, M.; Burgess, K., *J. Comb. Chem.* **2001**, *3*, 102-116.

58. Pattarawarapan, M.; Zaccaro, M. C.; Saragovi, U.; Burgess, K., *J. Med. Chem.* **2002**, *45*, 4387-4390.

59. Lee, H. B.; Pattarawarapan, M.; Roy, S.; Burgess, K., *Chem. Commun.* **2003**, 1674-1675.

60. Lee, H. B.; Zaccaro, M. C.; Pattarawarapan, M.; Roy, S.; Saragovi, H. U.; Burgess, K., *J. Org. Chem.* **2004**, *69*, 701-713.

61. Zaccaro, M. C.; Lee, H. B.; Pattarawarapan, M.; Xia, Z.; Caron, A.; Burgess, K.; Saragovi, H. U., *Chem. & Biol.* **2005**, *12*, 1015-1028.

62. Maliartchouk, S.; Debeir, T.; Beglova, N.; Cuello, A.; Gehring, K.; Saragovi, H., *J. Biol. Chem.* **2000**, *275*, 9946-9956.

63. Shi, Z.; Birman, E.; Saragovi, H. U., *Developmental Neurobiology* **2007**, *67*, 884-894.

64. Bruno, M. A.; Clarke, P. B. S.; Seltzer, A.; Quirion, R.; Burgess, K.; Cuello, A. C.; Saragovi, H. U., *J. Neuroscience* **2004**, *24*, 8009-8018.

65. Xie, Y.; Tisi, M. A.; Yeo, T. T.; Longo, F. M., *J. Biol. Chem.* **2000**, *275*, 29868-29874.

66. Li, B.; Tom, J. Y. K.; Oare, D.; Yen, R.; Fairbrother, W. J.; Wells, J. A.; Cunningham, B. C., *Science* **1995**, *270*, 1657-1660.

67. Wells, J. A., *Science* **1996**, *273*, 449-450.

68. Feng, Y.; Burgess, K., *Chem. Eur. J.* **1999**, *5*, 3261-3272.

69. Pattarawarapan, M.; Reyes, S.; Xia, Z.; Zaccaro, M. C.; Saragovi, H. U.; Burgess, K., *J. Med. Chem.* **2003**, *46*, 3565-3567.

70. LeSauteur, L.; Cheung, N. K. V.; Lisbona, R.; Saragovi, H. U., *Nature Biotech.* **1996**, *14*, 1120-1122.

71. Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P., *Journal of Computational Chemistry* **2003**, *24*, 1999-2012.

72. Jakalian, A.; Jack, D. B.; Bayly, C. I., *Journal of Computational Chemistry* **2002**, *23*, 1623-1641.

73. Wiesmann, C.; de Vos, A. M., *Cell. Mol. Life Sci.* **2001**, *58*, 748-759.

74. Wiesmann, C.; Ultsch, M. H.; Bass, S. H.; de Vos, A. M., *Nature* **1999**, *401*, 184-188.

75. Naim, M.; Bhat, S.; Rankin Kathryn, N.; Dennis, S.; Chowdhury Shafinaz, F.; Siddiqi, I.; Drabik, P.; Sulea, T.; Bayly Christopher, I.; Jakalian, A.; Purisima Enrico, O., *Journal of Chemical Information and Modeling* **2007**, *47*, 122-133.

76. Bopp, F., *Liebigs Ann. Chem.* **1849**, *69*, 16-37.

77. Kohler, A., *Liebigs Ann. Chem.* **1865**, *134*, 367-372.

78. Curties, T., *Chem. Ber.* **1883**, *16*, 753-757.

79. Martins, M. B.; Carvalho, I., *Tetrahedron* **2007**, *63*, 9923-9932.

80. Fischer, P. M., *Journal of Peptide Science* **2003**, *9*, 9-35.

81. Krchnak, V.; Weichsel, A. S.; Cabel, D.; Flegelova, Z.; Lebl, M., *Molecular Diversity* **1996**, *1*, 149-164.

82. McCleland, K.; Milne, P. J.; Lucieto, F. R.; Frost, C.; Brauns, S. C.; van de Venter, M.; Du Plessis, J.; Dyason, K., *Journal of Pharmacy and Pharmacology* **2004**, *56*, 1143-1153.

83. Davies, J. S., *Amino Acids, Peptides, and Proteins* **2006**, *35*, 272-352.

84. Pons, J.-F.; Fauchere, J.-L.; Lamaty, F.; Molla, A.; Lazaro, R., *Eur. J. Org. Chem.* **1998**, 853-859.

85. Royo, M.; Van Den Nest, W.; del Fresno, M.; Frieden, A.; Yahalom, D.; Rosenblatt, M.; Chorev, M.; Albericio, F., *Tetrahedron Lett.* **2001**, *42*, 7387-7391.

86. Li, W.-R.; Yang, J. H., *J. Comb. Chem.* **2002**, *4*, 106-108.

87. Wennemers, H., *Chimia* **2003**, *57*, 237-240.

88. Ressurreicao, A. S. M.; Bordessa, A.; Civera, M.; Belvisi, L.; Gennari, C.; Piarulli, U., *J. Org. Chem.* **2008**, *73*, 652-660.

89. Long, D. D.; Tennant-Eyles, R. J.; Estevez, J. C.; Wormald, M. R.; Dwek, R. A.; Smith, M. D.; Fleet, G. W. J., *Journal of the Chemical Society, Perkin Transactions 1* **2001**, 807-813.

90. Folkes, A.; Roe, M. B.; Sohal, S.; Golec, J.; Faint, R.; Brooks, T.; Charlton, P., *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2589-2592.

91. Wang, S.; Golec, J.; Miller, W.; Milutinovic, S.; Folkes, A.; Williams, S.; Brooks, T.; Hardman, K.; Charlton, P.; Wren, S.; Spencer, J., *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2367-2370.

92. Brooks, T. D.; Wang, S. W.; Bruenner, N.; Charlton, P. A., *Anti-Cancer Drugs* **2004**, *15*, 37-44.

93. Einholm, A. P.; Pedersen, K. E.; Wind, T.; Kulig, P.; Overgaard, M. T.; Jensen, J. K.; Bodker, J. S.; Christensen, A.; Charlton, P.; Andreasen, P. A., *Biochem. J.* **2003**, *373*, 723-732.

94. Acharya, A. N.; Ostresh, J. M.; Houghten, R. A., *J. Comb. Chem.* **2001**, *3*, 612-623.

95. Bianco, A.; Sonksen, C. P.; Roepstorff, P.; Briand, J.-P., *J. Org. Chem.* **2000**, *65*, 2179-2187.

96. Baures, P. W.; Ojala, W. H.; Costain, W. J.; Ott, M. C.; Gleason, W. B.; Mishra, R. K.; Johnson, R. L., *J. Med. Chem.* **1997**, *40*, 3594-3600.

97. Golebiowski, A.; Klopfenstein, S. R.; Chen, J. J.; Shao, X., *Tetrahedron Lett.* **2000**, *41*, 4841-4844.

98. Wang, D.-X.; Liang, M.-T.; Tian, G.-J.; Lin, H.; Liu, H.-Q., *Tetrahedron Lett.* **2002**, *43*, 865-867.

99. Horton, D. A.; Bourne, G. T.; Smythe, M. L., *Molecular Diversity* **2002**, *5*, 289-304.

100. Hulme, C.; Gore, V., *Current Medicinal Chemistry* **2003**, *10*, 51-80.

101. Cho, S.; Keum, G.; Kang, S. B.; Han, S.-Y.; Kim, Y., *Molecular Diversity* **2003**, *6*, 283-286.

102. Sollis, S. L., *J. Org. Chem.* **2005**, *70*, 4735-4740.

103. Chucholowski, A.; Masquelin, T.; Obrecht, D.; Stadlwieser, J.; Villalgordo, J. M., *Chimia* **1996**, *50*, 525-530.

104. Kennedy, A. L.; Fryer, A. M.; Josey, J. A., *Org. Lett.* **2002**, *4*, 1167-1170.

105. Santagada, V.; Fiorino, F.; Perissutti, E.; Severino, B.; Terracciano, S.; Cirino, G.; Caliendo, G., *Tetrahedron Lett.* **2003**, *44*, 1145-1148.

106. Kilian, G.; Jamie, H.; Brauns, S. C. A.; Dyason, K.; Milne, P. J., *Pharmazie* **2005**, *60*, 305-309.

107. Giraud, M.; Bernad, N.; Martinez, J.; Cavelier, F., *Tetrahedron Lett.* **2001**, *42*, 1895-1897.

108. Mehta, A.; Jaouhari, R.; Benson, T. J.; Douglas, K. T., *Tetrahedron Lett.* **1992**, *33*, 5441-5444.

109. Williams, R. M.; Glinka, T.; Kwast, E.; Coffman, H.; Stille, J. K., *J. Am. Chem. Soc.* **1990**, *112*, 808-821.

110. Nicolaou, K. C.; Estrada, A. A.; Zak, M.; Lee, S. H.; Safina, B. S., *Angew. Chem., Int. Ed.* **2005**, *44*, 1378-1382.

111. Endo, Y.; Hirano, M.; Driedger, P. E.; Stabel, S.; Shudo, K., *Bioorg. Med. Chem. Lett.* **1997**, *7*, 2997-3000.

112. Zaccaro, M. C.; B.Lee, H.; Pattarawarapan, M.; Xia, Z.; Caron, A.; L'Heureux., P.-J.; Bengio, Y.; Burgess, K.; Saragovi, H. U., *Chemistry & Biol.* **2005**, *12*, 1015-1028.

113. Zaccaro, M. C.; Lee, H. B.; Pattarawarapan, M.; Xia, Z.; Caron, A.; Burgess, K.; Saragovi, H. U., *Nat. Biotech.* **2003**, submitted.

114. Wipf, P., *Chem. Rev.* **1995**, *95*, 2115-2134.

115. Roy, R. S.; Gehring, A. M.; Milne, J. C.; Belshaw, P. J.; Walsh, C. T., *Natural Product Reports* **1999**, *16*, 249-263.

116. Yeh, V. S. C., *Tetrahedron* **2004**, *60*, 11995-12042.

117. Riego, E.; Hernandez, D.; Albericio, F.; Alvarez, M., *Synthesis* **2005**, 1907-1922.

118. Jin, Z., *Natural Product Reports* **2006**, *23*, 464-496.

119. Wagner, B.; Schumann, D.; Linne, U.; Koert, U.; Marahiel, M. A., *J. Am. Chem. Soc.* **2006**, *128*, 10513-10520.

120. Riedrich, M.; Harkal, S.; Arndt, H.-D., *Angew. Chem., Int. Ed.* **2007**, *46*, 2701-2703.

121. Moisan, L.; Odermatt, S.; Gombosuren, N.; Carella, A.; Rebek, J., Jr., *Eur. J. Org. Chem.* **2008**, 1673-1676.

122. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J., *Adv. Drug Delivery Rev.* **1997**, *23*, 3-25.

123. Weaver, G. W., *Science of Synthesis* **2004**, *13*, 219-251.

124. Stolle, R., *Berichte der Deutschen Chemischen Gesellschaft* **1899**, *32*, 797-798.

125. Pellizzari, G., *Real. Accad. dei Lincei* **1899**, *8*, 327-332.

126. Hetzheim, A.; Moeckel, K., *Advances in Heterocyclic Chemistry* **1966**, *7*, 183-224.

127. Hill, J., *Comprehensive Heterocyclic Chemistry II* **1996**, *4*, 267-287, 905-1006.

128. Roda, K. P.; Vansdadia, R. N.; Parekh, H., *Journal of the Indian Chemical Society* **1988**, *65*, 807-809.

129. Raman, K.; Parmar, S. S.; Salzman, S. K., *J. Pharm. Sci.* **1989**, *78*, 999-1002.

130. Ergenc, N.; Buyuktimkin, S.; Capan, G.; Baktir, G.; Rollas, S., *Pharmazie* **1991**, *46*, 290-291.

131. Saxena, V. K.; Singh, A. R.; Agarwai, R. K.; Mehra, S. C., *Journal of the Indian Chemical Society* **1983**, *60*, 575-577.

132. Mazzone, G.; Bonina, F.; Puglisi, G.; Panico, A. M.; Arrigo Reina, R., *Farmaco, Edizione Scientifica* **1984**, *39*, 414-420.

133. Moutevelis-Minakakis, P.; Photaki, I., *Journal of the Chemical Society, Perkin Transactions 1: Organic and Bio-Organic Chemistry (1972-1999)* **1985**, 2277-2281.

134. Zhang, X.; Breslav, M.; Grimm, J.; Guan, K.; Huang, A.; Liu, F.; Maryanoff, C. A.; Palmer, D.; Patel, M.; Qian, Y.; Shaw, C.; Sorgi, K.; Stefanick, S.; Xu, D., *J. Org. Chem.* **2002**, *67*, 9471-9474.

135. Davies, J. R.; Kane, P. D.; Moody, C. J., *Tetrahedron* **2004**, *60*, 3967-3977.

136. Wipf, P.; Miller, C. P., *J. Org. Chem.* **1993**, *58*, 3604-3606.

137. Paquette, L. A.; Mitzel, T. M.; Isaac, M. B.; Crasto, C. F.; Schomer, W. W., *J. Org. Chem.* **1997**, *62*, 8960.

138. Icke, R. N.; Wisegarver, B. B.; Alles, G. A., *Org. Synth.* **1945**, *25*, 89-91.

139. Lin, C.-C.; Wu, J.-C.; Chang, T.-T.; Chang, W.-Y.; Yu, M.-L.; Tam, A. W.; Wang, S.-C.; Huang, Y.-H.; Chang, F.-Y.; Lee, S.-D., *J. Clin. Microbiol.* **2000**, *38*, 3915-3918.

140. Lu, B.; Smyth, M. R.; O'Kennedy, R., *Analyst* **1996**, *121*, 29R-32R.

141. Yarmuch, M. L.; Weiss, A. M.; Antonsen, K. P.; Odde, D. J.; Yarmush, D. M., *Biotech. Adv.* **1992**, *10*, 413-446.

142. Askari, M.; Alarie, J. P.; Moreno-Bondi, M.; Vo-Dinh, T., *Biotechnol. Prog.* **2001**, *17*, 543-552.

143. Stokes, D. L.; Griffin, G. D.; Vo-Dinh, T., *Fresenius J. Anal. Chem.* **2001**, *369*, 295-301.

144. Mendoza, L. G.; McQuary, P.; Mongan, A.; Gangadharan, R.; Brignac, S.; Eggers, M., *BioTechniques* **1999**, *27*, 778-788.

145. Banerjee, A.; Srilatha, N. S.; Murthy, G. S., *Biochim. Biophys. Acta* **2002**, *1569*, 21-30.

146. van Dijk, M. A.; van de Winkel, J. G. J., *Curr. Opin. Chem. Biol.* **2001**, *5*, 368-374.

147. Cragg, M. S.; French, R. R.; Glennie, M. J., *Curr. Opin. Immunol.* **1999**, *11*, 541-547.

148. Vaswani, S. K.; Hamilton, R. G., *Annals Allergy, Asthma, & Immunol.* **1998**, *81*, 105-119.

149. Bond, A.; Jones, M. G.; Hay, F. C., *J. Immunol. Meth.* **1993**, *166*, 27-33.

150. Boden, V.; Winzerling, J. J.; Vijayalakshmi, M.; Porath, J., *J. Immunol. Meth.* **1995**, *181*, 225-232.

151. Huse, K.; Bohme, H.-J.; Scholz, G. H., *J. Biochem. Biophys. Methods* **2002**, *51*, 217-231.

152. Labrou, N.; Clonis, Y. D., *J. Biotech.* **1994**, *36*, 95-119.

153. Verdoliva, A.; Cassani, G.; Fassina, G., *J. Chrom. B: Biomed. Appl.* **1995**, *664*, 175-183.

154. Labrou, N. E., *J. of Chromatography B* **2003**, *790*, 67-78.

155. Langone, J. J., *Adv. Immunol.* **1982**, *32*, 157-252.

156. Kihira, Y.; Aiba, S., *J. Chrom.* **1992**, *597*, 277-283.

157. Deisenhofer, J., *Biochem.* **1981**, *20*, 2361-2370.

158. Saha, K.; Bender, F.; Gizeli, E., *Anal. Chem.* **2003**, *75*, 835-842.

159. Langone, J. J., *J. Immunol. Meth.* **1982**, *55*, 277-296.

160. Hale, G.; Drumm, A.; Harrison, P.; Phillips, J., *J. Immunol. Meth.* **1994**, *171*, 15-21.

161. Guss, B.; Eliasson, M.; Olsson, A.; Uhlen, M.; Frej, A.-K.; Jornvall, H.; Flock, J. I.; Lindberg, M., *EMBO* **1986**, *5*, 1567-1575.

162. Bjorck, L.; Kronvall, G., *J. Immunol.* **1984**, *133*, 969-974.

163. Akerstrom, B.; Brodin, T.; Reis, K.; Bjorck, L., *J. Immunol.* **1985**, *135*, 2589-2592.

164. Godfrey, M. A. J.; Kwasowski, P.; Clift, R.; Marks, V., *J. Immunol. Meth.* **1992**, *149*, 21-27.

165. Fuglistaller, P., *J. Immunol. Meth.* **1989**, *124*, 171-177.

166. Fassina, G.; Verdoliva, A.; Odierna, M. R.; Ruvo, M.; Cassini, G., *J. Mol. Recog.* **1996**, *9*, 564-569.

167. Guerrier, L.; Flayeux, I.; Schwarz, A.; Fassina, G.; Boschetti, E., *J. Mol. Recog.* **1998**, *11*, 107-109.

168. Palombo, G.; De Falco, S.; Tortora, M.; Cassani, G.; Fassina, G., *J. Mol. Recog.* **1998**, *11*, 243-246.

169. Fassina, G.; Ruvo, M.; Palombo, G.; Verdoliva, A.; Marino, M., *J. Biochem. Biophys. Methods* **2001**, *49*, 481-490.

170. Teng, S. F.; Sproule, K.; Hussain, A.; Lowe, C. R., *J. Mol. Recog.* **1999**, *12*, 67-75.

171. Teng, S. F.; Sproule, K.; Husain, A.; Lowe, C. R., *J. Chromatogr. B* **2000**, *740*, 1-15.

172. Lowe, C. R.; Burton, S. J.; Burton, N. P.; Alderton, W. K.; Pitts, J. M.; Thomas, J. A., *Tibtech* **1992**, *10*, 442-448.

173. Sproule, K.; Morrill, P.; Pearson, J. C.; Burton, S. J.; Hejnaes, K. R.; Valore, H.; Ludvigsen, S.; Lowe, C. R., *J. Chromatogr. B* **2000**, *740*, 17-33.

174. Roque, A. C. A.; Taipa, M. A.; Lowe, C. R., *J. Mol. Recog.* **2005**, *18*, 213-224.

175. Roque, A. C. A.; Taipa, M. A.; Lowe, C. R., *J. Chromatogr. A* **2005**, *1064*, 157-167.

176. Wang, Y.; Li, T., *Biotechnol. Prog.* **2002**, *18*, 524-529.

177. Kabir, S., *Immunol. Invest.* **2002**, *31*, 263-278.

178. Roque, A. C. A.; Silva, C. S. O.; Taipa, M. A., *J. Chromatogr. A* **2007**, *1160*, 44-55.

179. Palombo, G.; Rossi, M.; Cassani, G.; Fassina, G., *J. Mol. Recog.* **1998**, *11*, 247-249.

180. Newcombe, A. R.; Cresswell, C.; Davies, S.; Watson, K.; Harris, G.; O'Donovan, K.; Francis, R., *J. Chromatogr. B* **2005**, *814*, 209-215.

181. Chhatre, S.; Titchener-Hooker, N. J.; Newcombe, A. R.; Keshavarz-Moore, E., *Nat. Protocols* **2007**, *2*, 1763-1769.

182. Gouda, H.; Torigoe, H.; Saito, A.; Sato, M.; Arata, Y.; Shimada, I., *Biochem.* **1992**, *31*, 9665-9672.

183. Gouda, H.; Shiraishi, M.; Takahashi, H.; Kato, K.; Torigoe, H.; Arata, Y.; Shimada, I., *Biochem.* **1998**, *37*, 129-136.

184. Cedergren, L.; Andersson, R.; Jansson, B.; Uhlen, M.; Milsson, B., *Protein Engineering* **1993**, *6*, 441-448.

185. Li, R.; Dowd, V.; Stewart, D. J.; Burton, S. J.; Lowe, C. R., *Nat. Biotech.* **1998**, *16*, 190-195.

186. O'Connor, S. D.; Smith, P. E.; Al-Obeidi, F.; Pettitt, B. M., *J. Med. Chem.* **1992**, *35*, 2870-2881.

187. Pettitt, B. M.; Matsunaga, T.; Al-Obeidi, F.; Gehrig, C.; Hruby, V. J.; Karplus, M., *Biophy. J. Biophys. Soc.* **1991**, *60*, 1540-1544.

188. Campbell, J. E.; Englund, E. E.; Burke, S. D., *Org. Lett.* **2002**, *4*, 2273-2275.

189. Carpino, L. A.; El-Faham, A., *J. Am. Chem. Soc.* **1995**, *117*, 5401-5402.

190. Chiou, A.-J.; Ong, G.-T.; Wang, K.-T.; Chiou, S.-H.; Wu, S.-H., *Biochem. Biophys. Res. Commun.* **1996**, *219*, 572-579.

191. Wels, B.; Kruijtzer, J. A. W.; Liskamp, R. M. J., *Org. Lett.* **2002**, *4*, 2173-2176.

192. Achari, A.; Hale, S. P.; Howard, A. J.; Clore, G. M.; Gronenborn, A. M.; Hardman, K. D.; Whitlow, M., *Biochem.* **1992**, *31*, 10449-10457.

193. Derrick, J. P.; Wigley, D. B., *J. Mol. Biol.* **1994**, *243*, 906-918.

194. Gallagher, T.; Alexander, P.; Bryan, P.; Gilliland, G. L., *Biochem.* **1994**, *33*, 4721-4729.

195. Sauer-Eriksson, A. E.; Kleywegt, G. J.; Uhlen, M.; Jones, T. A., *Structure* **1995**, *3*, 265-278.

196. Bogan, A. A.; Thorn, K. S., *J. Mol. Biol.* **1998**, *280*, 1-9.

197. Fantl, W. J.; Johnson, D. E.; Williams, L. T., *Annu. Rev. Biochem* **1993**, *62*, 453-481.

198. Li, Y.; Dias, J. R., *Chem. Rev.* **1997**, *97*, 283-304.

199. Klemm, J. D.; Schreiber, S. L.; Crabtree, G. R., *Annu. Rev. Immunol.* **1998**, *16*, 569-592.

200. Zutshi, R.; Brickner, M.; Chemielewski, J., *Cur. Opin. Chem. Bio.* **1998**, *2*, 62-66.

201. Mammen, M.; Choi, S.-K.; Whitesides, G. M., *Angew. Chem., Int. Ed.* **1998**, *37*, 2754-2794.

202. Kim, R. M.; Manna, M.; Hutchins, S. M.; Griffin, P. R.; Yates, N. A.; Bernick, A. M.; Chapman, K. T., *Proc. Natl. Acad. Sci.* **1996**, *93*, 10012-10017.

203. Qureshi, S. A.; Kim, R. M.; Konteatis, Z.; Biazzo, D. E.; Motamedi, H.; Rodrigues, R.; Boice, J. A.; Calaycay, J. R.; Bednarek, M. A.; Griffin, P.; Gao, Y.-D.; Chapman, K.; Mark, D. F., *Proc. Natl. Acad. Sci.* **1999**, *96*, 12156-12161.

204. Zelikin, A.; Shastri, V. R.; Langer, R., *J. Org. Chem* **1999**, *64*, 3379-3380.

205. Potapova, I. A.; Purygin, P. P.; Lipatov, I. S.; Belousova, Z. P.; Yakimova, N. A.; Tezikov, Y. V.; Selezneva, E. S., *Pharm. Chem. J.* **2001**, *35*, 588-590.

206. Cardona, C. M.; Jannach, S. H.; Huang, H.; Itojima, Y.; Leblanc, R. M.; Gawley, R. E., *Helv. Chim. Acta* **2002**, *85*, 3532-3558.

207. Dressman, B. A.; Spangle, L. A.; Kaldor, S. W., *Tetrahedron Lett.* **1996**, *37*, 937-940.

APPENDIX A

EXPERIMENTAL FOR CHAPTER II

General Methods. All the amino acids used had the L-configuration and they were purchased from NovaBiochem, Advance BioTech or Chem-Impex. All chemicals were obtained from commercial suppliers and used without further purification. *N*-Hydroxybenzotriazole (HOBt), *N*-methylmorpholine, *N,N*-diisopropylethylamine were purchased from Aldrich. Ethyl-(*N*',*N*'-dimethylamino)propylcarbodiimide (EDC) was obtained from Advanced ChemTech. Reverse phase high performance liquid chromatography (RP-HPLC) was carried out on Vydac C-18 column with the dimension of 25 x 0.46 cm for analytical studies, and on XTerra C-18 column with the dimension of 10 x 1.9 cm for preparative work. All HPLC experiments were performed using gradient conditions. The eluents used were: solvent A (H_2O with 0.1% TFA) and solvent B (CH_3CN with 0.1% TFA). The flow rate used was 1.0 mL/min for analytical, and 10 mL/min for preparative HPLC.

General Procedure for Preparation of 1a-i. The Rink Amide resin MBHA (1.0 g; 0.72 mmol) was treated with 20% piperidine in DMF (2 x 10 mL, 10 min), and washed. The resin was treated with FMOC-Hse-(Trt)-OH (2 equiv), HBTU (1.95 equiv), HOBt (2 equiv), and DIEA (3 equiv) in DMF (6 mL) for 24 h. After washing and FMOC deprotection, the resin was then treated with FMOC-Lys(Boc)-OH (2 equiv), HBTU (1.95 equiv), HOBt (2 equiv), and DIEA (3 equiv) in DMF (6 mL) for 24 h. After washing and FMOC deprotection, the resin was then treated with FMOC-Glu(OtBu)-OH (2 equiv), HBTU (1.95 equiv), HOBt (2 equiv), and DIEA (3 equiv) in DMF (6 mL) for 24 h. The washing cycle and FMOC deprotection were repeated, and 2-fluoro-5-nitrobenzoic acid moiety was introduced by treating the resin with 2-fluoro-5-nitrobenzoic acid (2 equiv), HBTU (1.95 equiv), HOBt (2 equiv), and 2,6-lutidine (15 equiv) in CH_2Cl_2/DMF (1:1, 6 mL) for 12 h. The side chain protecting group (Trt) of the homoserine was removed by treatment with 1% TFA and 5% HSi^iPr_3 in CH_2Cl_2 (3×2 min, or until color disappeared). After the resin was washed, the macrocyclization step

was carried out by treating the supported peptide with K_2CO_3 (10 equiv) in DMF at 25°C. After gentle shaking for 24 h, the peptide-resin was washed then dried in vacuo for 4 h. Resin containing side chain protected cyclic peptidomimetic D3 with the nitro group was swelled in CH_2Cl_2 in a fritted syringe for 30 min and treated with 2M $SnCl_2 \bullet 2H_2O$ in DMF for 12 h. After the washing and drying under vacuum for 2 h, the resulting arylamine was treated with the short, medium or long linker (4 equiv), PyBrOP (4 equiv), and 2,6-lutidine (10 equiv) in DMF for 24 h. The resin was then washed, dried, and split into two portions. One portion was subjected to FMOC deprotection by 20% piperidine in DMF (2 x 10 mL, 10 min). After that the peptide was cleaved from the resin by treatment with a mixture of 90% TFA, 5% HSi^iPr3, and 5% H_2O for 3 h. The cleavage solution was separated from the resin by filtration. After most of the cleavage cocktail (about 90%) was evaporated in vacuo, the crude peptide was triturated using anhydrous ethyl ether, dissolved in H_2O, and then lyophilized to give the crude product. Preparative HPLC (Beckman System, 10-80% CH_3CN in H_2O + 0.1% TFA in 40 min) was carried out to give **D3** derivative peptidomimetics. Another portion of dry resin was swelled in CH_2Cl_2 for 30 min, washed with DMF, treated with 20% piperidine in DMF (2 x 10mL, 10 min), and washed. After the washing dichlorotriazinylaminofluorescein (DTAF, 2 equiv), DIEA (3 equiv) in 3:1 mixture of $CHCl_3$:DMSO were added and the mixture was shaken for 3 h. The resin was washed and subjected to the above cleavage conditions. The crude product was purified by preparative HPLC to give dye-labeled peptidomimetics.

Glass vials with caps (short form style, 1 Dr, VWR International products) were used as reaction vessels. Stock solutions of all reagents in DMSO were prepared: 0.030 M of first cleavage portion (non-tagged) and 0.033 M of second cleavage portion (dye-tagged). Solid K_2CO_3 (4 equiv) was first added to each vial, followed by equal volumes of the two stock solutions of the non-tagged and the dye-tagged peptidomimetic. The reaction vessels were capped and stirred at 25 °C for 24 h. After analysis of the crude purity on analytical HPLC, the solution in each vial was lyophilized to remove DMSO. The solid materials were re-dissolved in 1:1 mixture of H_2O/CH_3CN, and then purified by preparative HPLC to yield the final products **1a-i**.

Materials and Methods for Biological Assays. Cell Lines—NIH-TrkA cells are NIH3T3 fibroblasts transfected with human *trkA* cDNA. NIH-TrkC cells are NIH3T3 fibroblasts transfected with human *trkC* cDNA. NIH-IGF-1R cells are NIH3T3 fibroblasts transfected with human *IGF-1R* cDNA. All cells are stably transfected subclones that express high levels of their receptor. Cells are grown under drug selection (0.5 mg/ml G418) and are routinely screened for receptor expression by FACScan using monoclonal antibodies directed to the receptor extracellular domains.

Antibodies—mouse anti-TrkA mAb 5C3, mouse anti-TrkC mAb 2B7, mouse anti-p75 mAb MC912, and mouse anti-IGF-1R mAb alphaIR3 were purified using protein G-Sepharose (Pharmacia, Baie d'Urfe, Quebec, Canada). They have all been described.

FACScan analysis—Cells (2 x 105) in 100 μl of FACScan binding buffer (phosphate-buffered saline, 0.5% BSA, and 0.1% NaN_3) were immunostained as described. Saturating mAbs, or control non-binding mouse IgGs were added to cells for 20 min at 4 °C, excess primary antibody was washed off, and cells were immunostained with fluoresceinated goat anti-mouse IgG (FITC-G-α-M) secondary antibody. As cellular controls NIH-3T3 cells not expressing NGF receptors were used (e.g. NIH-IGF-1R). Cells were acquired on a FACScan, and bell-shaped histograms were analyzed using LYSIS II and the CellQuest-pro program as described.

FACScan analysis of peptidomimetics—Peptidomimetics labeled with FITC were used in FACScan binding studies, as described above. The assay was slightly modified to extend the incubation of the "primary" reagent to 40 minutes, followed by two washes to remove unbound material. There was no secondary reagent added, as the compounds are directly labeled.

FACScan binding competition assays—Blocking to the binding of FITC-labeled peptidomimetics were studied by pre-incubation of the cells with increasing concentrations of TrkA ligands anti-TrkA mAb 5C3, for 30 min at 4°C. Then, the FITC-labeled peptidomimetic was added, without previous washing, and the study progressed as described above.

Ligand Binding—^{125}I[NGF] (73.1 mCi/mg; NEN Life Science Products) binding assays and Scatchard plot analysis were done on cells as described. NIH-TrkA cells

(1x106 per point) were added to serial dilutions of ^{125}I[NGF] in BB at 4°C, in the presence or absence of competitor (cold NGF or peptidomimetic). NIH-3T3 wild type cells not expressing NGF receptors were used to assess nonspecific background (<15% of total binding).

Proliferation and Survival Assays—Cells (5,000–10,000 cells/well) were added to 96-well plates (Becton Dickinson, Lincoln Park, NJ) and cultured either in media containing 5% fetal bovine serum or in serum free media with 0.1% BSA (SFM). Ligands consisted of serial dilutions of neurotrophins or peptidomimetics were then added. FITC-tagged and TEG-tagged mimetics were tested, but only the data for TEG-tagged mimetics are shown. A suboptimal dose of neurotrophins (0.1 – 0.2 nM, affording ~25% of survival) was used to test the effect of combination of NGF with peptidomimetics. The proliferative/survival profile of the cells was quantitated using the tetrazolium salt reagent (3-(4,5- dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, Sigma) and optical density (OD) readings as described. Assays were done 4-7 times, each assay n 4–8.

Receptor Dimerization Assays—Receptor cross-linking assays were carried out as reported elsewhere. Here, NIH-TrkA cells (106 cells/ml each group) were washed and placed in suspension in phosphate-buffered saline (pH 7.5) at 4°C. Then they were exposed to the indicated ligands (untreated control, NGF 10 nM, **1-ss-F*** (20 μM), or **1-ss-TEG** (20 μM)) for 30 min at 4°C. Following washing, cells were chemically cross-linked with 1 mM final disuccinimidyl suberate (DSS, Pierce) for 7 min. To quench unreacted DSS a 5-molar excess of ammonium acetate was added and after 1 min the cells were washed two times with HBSS at 4°C. Then each cell pellet was detergent solubilized (1% NP40 containing protease inhibitors), lysates were centrifuged (15,000 G, 15 min 4°C) to remove nuclei and detergent-insoluble material. Protein concentration for each of the clear lysates was determined using a detergent-compatible BioRad kit. Equal protein (20 μg/lane) of each sample were resolved by SDS-PAGE, and after Western transfer the membranes were analyzed by western blotting with anti-TrkA mAbs. Equal loading was further verified Coomassie blue staining of the gels.

Receptor Activation Assays—NIH-TrkA cells were cultured at 37°C in SFM for 12 hours to reduce their baseline tyrosine phosphorylation. Then the cells were exposed

to the indicated ligand (NGF, **1-ss**, untreated control) at 37°C for 20 minutes or for 2 hours. After washing in PBS, cells were detergent solubilized as above, and equal protein (20 μg/lane) of each of the clear lysates were resolved by SDS-PAGE, and after Western transfer the membranes were analyzed by western blotting with anti-pTyr mAb 4G10 (Promega), as previously reported; and total TrkA loading was verified using anti-TrkA mAb 5C3.

Methods for Docking/Modeling. 3-D structure of **1-ss**—The three dimensional structures of **1-ss** were built by using the BIOPOLYMER module of SYBYL. The initial structure was optimized by using Tripos force field. GAFF (Generalised amber force field) atom types were obtained using the antechamber module of AMBER. This facility automatically generates parameters that are compatible with the AMBER force field. Partial charges of **1-ss** were calculated by the AM1-BCC method.

Model of TrkA-D5 target for docking—The extracellular immunoglobulin-like domain (D5) of the TrkA receptor forms a high-affinity binding site for NGF. It is necessary and sufficient for NGF binding. The crystal structure of human TrkA-D5 in complex with human NGF with was retrieved from the Protein Data Bank (PDB code 1www). The bound NGF dimer and the crystallographic water molecules were removed from the pdb file. The program Reduce was used to add hydrogen atoms to the target atoms and to optimize the orientation of the polar hydrogens. Both N- and C- termini of the protein were modeled in the ionized state. Structure manipulation and visualization were done using SYBYL 7.3 (Tripos, Inc., St. Louis, MO, USA).

Docking of **1-ss**—To explore possible binding sites for **1-ss**, a rectangular box was constructed such that it almost completely enclosed the NGF dimer in the NGF-TrkA-D5 crystal complex. This box defined the volume to be explored in docking **1-ss** to TrkA-D5. A molecular docking program, WILMA, was used for docking. This program is a rigid body docker that takes as input a collection of low-energy conformers of the ligand, **1-ss**. OMEGA (OpenEye Scientific, Inc.) generated 200 conformers within the 20 kcal/mol cutoff. Each of these conformers was exhaustively docked within the rectangular search box by WILMA. The poses were ranked according to their electrostatic, hydrogen bonding, van der Waals and shape complementarity scores.

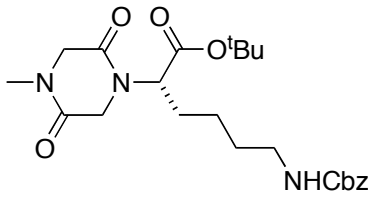
APPENDIX B

EXPERIMENTAL FOR CHAPTER III

General Methods. All the amino acids used had the L-configuration and they were purchased from NovaBiochem, Advance BioTech or Chem-Impex. All chemicals were obtained from commercial suppliers and used without further purification. *N*-Hydroxybenzotriazole (HOBt), *N*-methylmorpholine, *N,N*-diisopropylethylamine were purchased from Aldrich. Ethyl-(*N*',*N*'-dimethylamino)propylcarbodiimide (EDC) was obtained from Advanced ChemTech. Dichloromethane was obtained anhydrous by distillation over calcium hydride and THF was distilled over sodium metal and benzophenone. Reverse phase high performance liquid chromatography (RP-HPLC) was carried out on Vydac C-18 column with the dimension of 25 x 0.46 cm for analytical studies, and on Axia C-18 column with the dimension of 10 x 2.12 cm for preparative work. All HPLC experiments were performed using gradient conditions. The eluents used were: solvent A (H_2O with 0.1% TFA) and solvent B (CH_3CN with 0.1% TFA). The flow rate used was 1.0 mL/min for analytical, and 10 mL/min for preparative HPLC. NMR spectra of compounds **2-7** were recorded on Inova instruments at 300 MHz and at 500 MHz. NMR chemical shifts were expressed in ppm relative to internal solvent peaks, and coupling constants were measured in Hz.

General Procedure for Preparation of 2a-b. *N*-methyl Boc protected amino acid (1.0 equiv) and amino acid *tert*-butyl ester (1.0 equiv) was dissolved in dry CH_2Cl_2 (0.15 M). *N*-methylmorpholine (2.1 equiv) was added to the above solution followed with HOBt (1.1 equiv) and EDC (1.1 equiv). The resulting solution was stirred at room temperature for 12h. The reaction mixture was washed with H_2O (30 ml), 1N HCl (30 ml), saturated $NaHCO_3$ (30 ml) and brine (30 ml). The organic layer was separated and dried with Na_2SO_4 and concentrated to dryness to afford the crude product. Without further purification this dipeptide (1 equiv) was dissolved in a mixture of dry CH_2Cl_2 (32 equiv) and TFA (13 equiv) at 0 °C. It was allowed to warm up to 25 °C after 1h and stirred for another 1h until the reaction is complete. The reaction mixture was dried

under reduced pressure and the product was used directly for the next reaction. To a solution of crude secondary amine in CH_2Cl_2 (0.12 M) was added an aqueous solution of 0.5 M K_2CO_3 (2.1 equiv), and the mixture was cooled to 0 °C. To the vigorously stirred mixture was added bromoacetyl bromide (1.1 equiv) in one portion. After 2-4h (when no more starting material was present) the organic layer was separated and washed with H_2O (30 ml), 1N HCl (30 ml), saturated $NaHCO_3$ (30 ml) and brine (30 ml). The organic phase was separated and dried with Na_2SO_4 and concentrated to dryness. The product was used for the next reaction without further purification. The bromide was dissolved in a mixture of CH_2Cl_2 (0.10 M) and 50 % aqueous NaOH (v:v=10:1). The resulting mixture was stirred vigorously at 25 °C for 12 h until the starting material was consumed. The mixture was diluted with H_2O, and the organic layer was separated, washed with 1N HCl (30 ml) and brine (30 ml), dried with Na_2SO_4, and concentrated under reduced pressure. Flash chromatography of the residue provided pure **2a-b**.



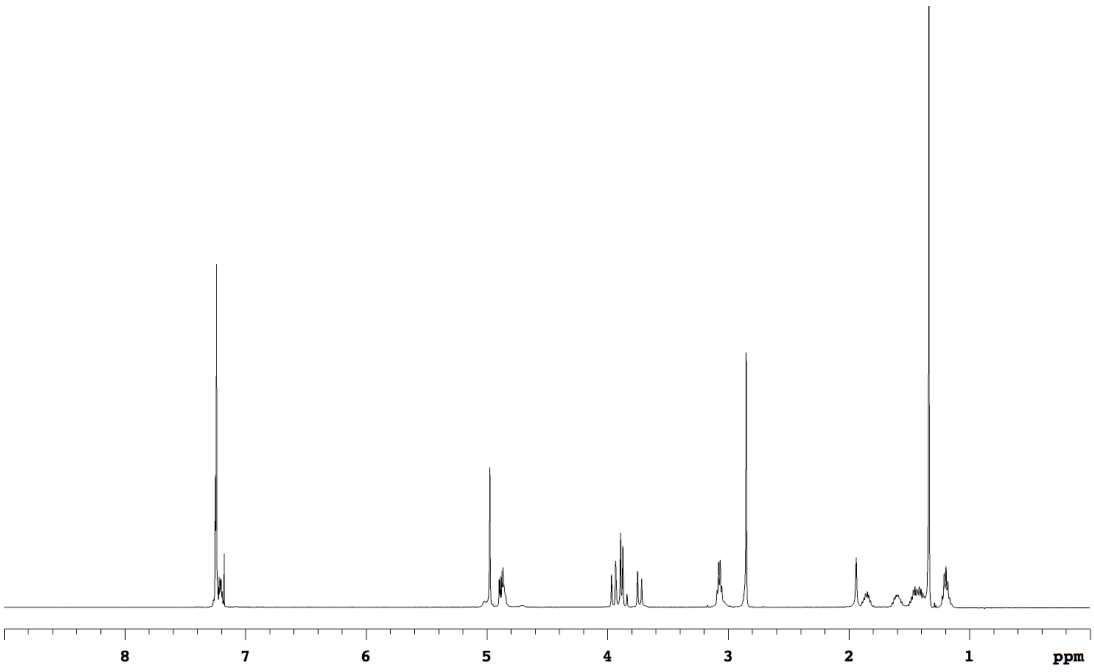
2a

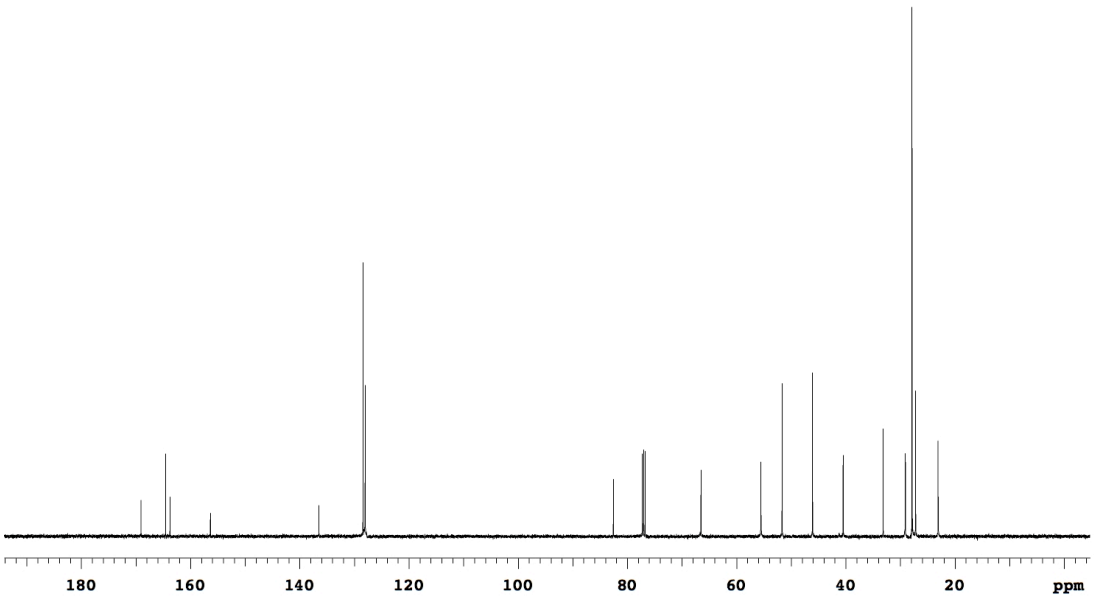
Compound **2a** was prepared from (0.58 g, 1.1 mmol) of Sar-Lys bromide derivative. Flash chromatography (2 % to 3 % MeOH/CH_2Cl_2) afforded 0.48 g (95 %) **2a** as a white solid.

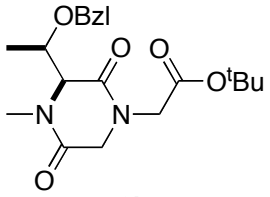
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.29-7.17 (m, 5H), 4.97 (s, 2H), 4.91-4.82 (m, 2H), 3.95 (d, 1H, J = 17.1 Hz), 3.91-3.85 (m, 2H), 3.73 (d, 1H, J = 17.1 Hz), 3.12-3.00 (m, 2H), 2.85 (s, 3H), 1.90-1.79 (m, 1H), 1.66-1.54 (m, 1H), 1.52-1.31 (m, 11H), 1.25-1.13 (m, 2H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 169.1, 164.5, 163.8, 156.4, 136.5, 128.4, 128.0, 128.0, 82.6, 66.5, 55.5, 51.7, 46.1, 40.5, 33.2, 29.1, 27.9, 27.2, 23.1

MS (APCI, m/z) 447.9 $(M+H)^+$


8
7
6
5
4
3
2
1
ppm


180
160
140
120
100
80
60
40
20
ppm

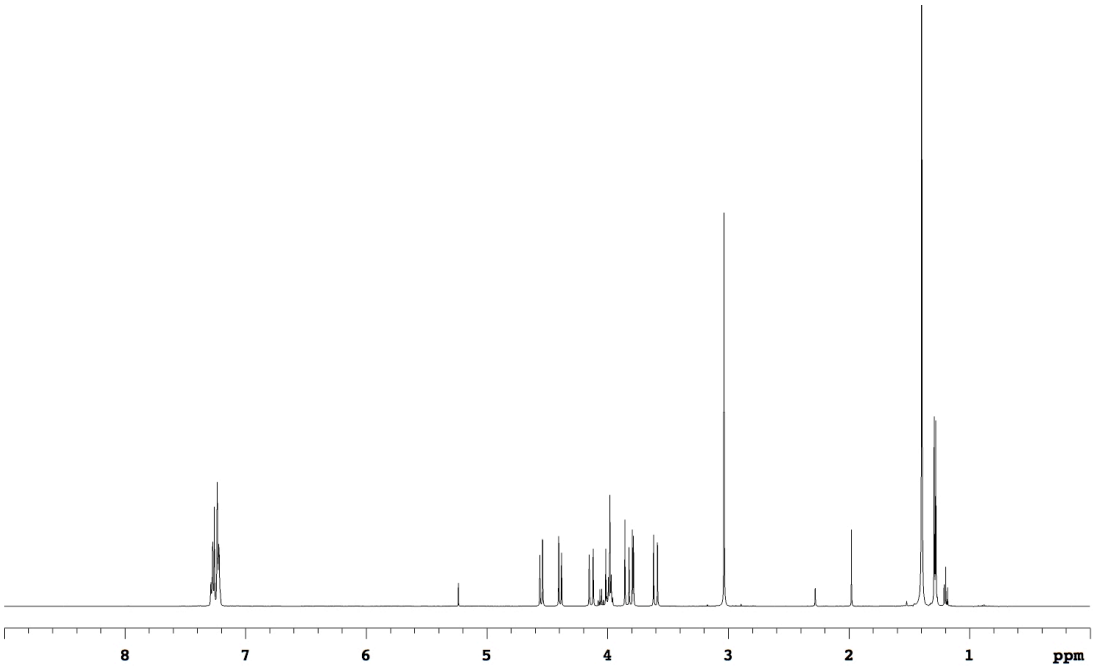


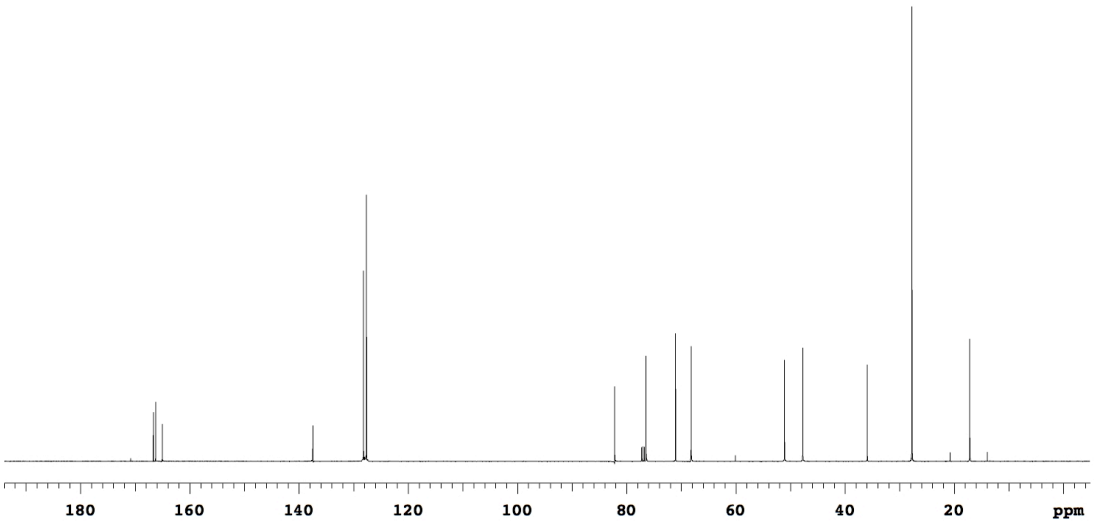
2b

Compound **2b** was prepared from (0.46 g, 1.0 mmol) of Thr-Gly bromide derivative. Flash chromatography (30 % to 70 % EtOAc/Hexanes) afforded 0.29 g (78 %) **2b** as a white solid.

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.31-7.19 (m, 5H), 4.55 (d, 1H, J = 11.5 Hz), 4.39 (d, 1H, J = 11.5 Hz), 4.13 (d, 1H, J = 16.4 Hz), 4.03-3.95 (m, 2H), 3.84 (d, 1H, J = 17.4 Hz), 3.79 (d, 1H, J = 4.6 Hz), 3.60 (d, 1H, J = 16.4 Hz), 3.03 (s, 3H), 1.40 (s, 9H), 1.29 (d, 3H, J = 6.3 Hz)

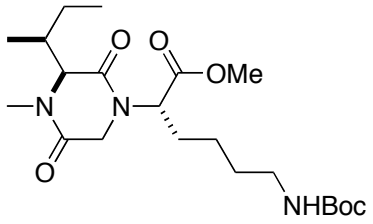
^{13}C NMR (125 MHz, $CDCl_3$) δ 166.7, 166.3, 165.1, 137.5, 128.2, 127.6, 127.6, 82.2, 76.5, 71.0, 68.2, 51.1, 47.7, 35.9, 27.7, 17.2





General Procedure for Preparation of 2c-d. *N*-methyl Boc protected amino acid (1.0 equiv) and amino acid methyl ester (1.0 equiv) was dissolved in dry CH_2Cl_2 (0.15 M). *N*-methylmorpholine (2.1 equiv) was added to the above solution followed with HOBt (1.1 equiv) and EDC (1.1 equiv). The resulting solution was stirred at room temperature for 12h. The resulting mixture was washed with H_2O (30 ml), 1N HCl (30 ml), saturated $NaHCO_3$ (30 ml) and brine (30 ml). The organic layer was separated and dried with Na_2SO_4 and concentrated to dryness to afford the crude product. Without further purification, a solution of this dipeptide in dry methanol (0.10 M) was stirred vigorously with 10 % Pd-C (0.1 equiv) in a H_2 atmosphere for 16 h at 25 °C until the reaction is complete. Filtration and evaporation of the solvent *in vacuo* yielded the crude product and it was used directly for the next reaction. To a solution of crude secondary amine in CH_2Cl_2 (0.12 M) was added an aqueous solution of 0.5 M K_2CO_3 (2.1 equiv), and the mixture was cooled to 0 °C. To the vigorously stirred mixture was added bromoacetyl bromide (1.1 equiv) in one portion. After 2-4h (when no more starting material was present) the organic layer was separated and washed with H_2O (30 ml), 1N HCl (30 ml), saturated $NaHCO_3$ (30 ml) and brine (30 ml). The organic phase was

separated and dried with Na_2SO_4 and concentrated to dryness. The product was used for the next reaction without further purification. The bromide was dissolved in a mixture of CH_2Cl_2 (0.10 M) and 50 % aqueous NaOH (v:v=10:1). The resulting mixture was stirred vigorously at 25 °C for 12 h until the starting material was consumed. The mixture was diluted with H_2O, and the organic layer was separated, washed with 1N HCl (30 ml) and brine (30 ml), dried with Na_2SO_4, and concentrated under reduced pressure. Flash chromatography with MeOH/CH_2Cl_2 mixtures provided pure **2c**.

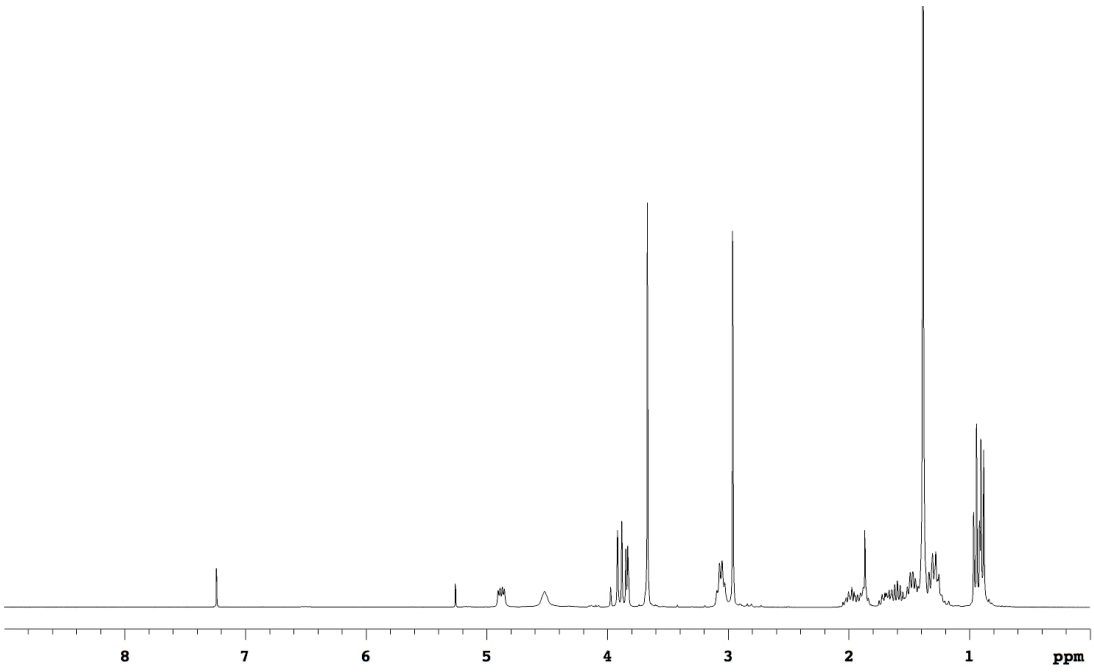


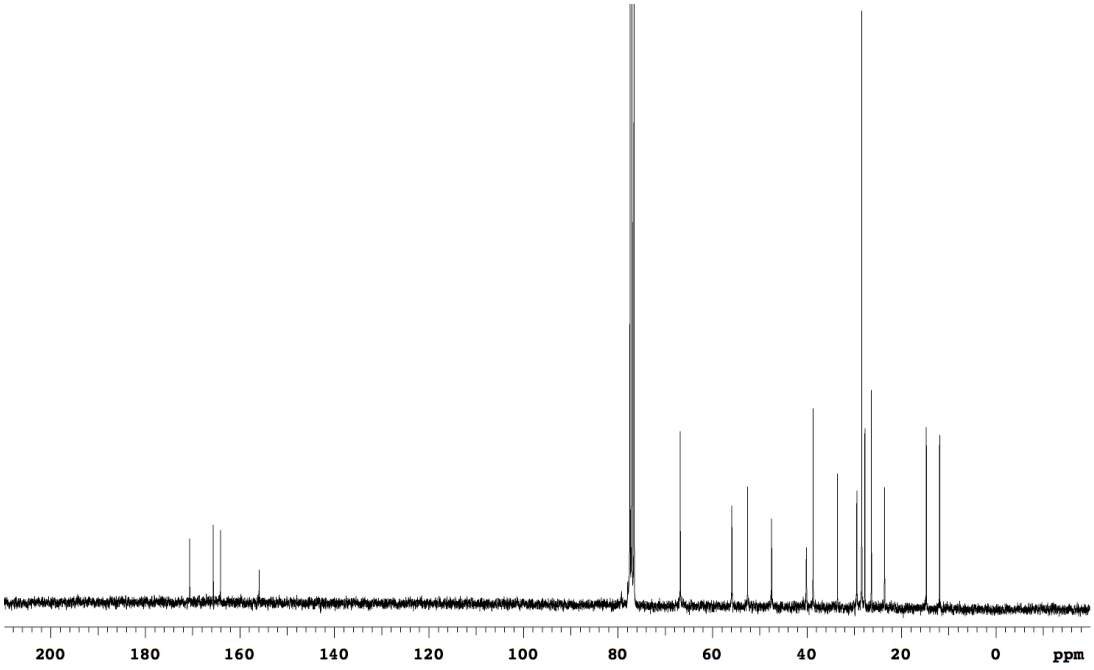
2c

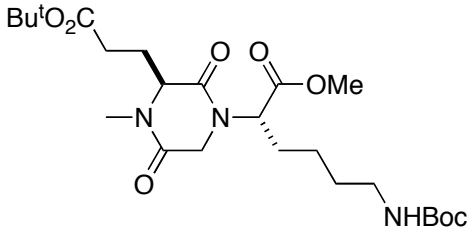
Compound **2c** was prepared from (1.40 g, 2.7 mmol) of Ile-Lys bromide derivative. Flash chromatography (2 % to 10 % MeOH/CH_2Cl_2) afforded 0.85 g (72 %) **2c** as a white solid.

^{1}H NMR (300 MHz, $CDCl_3$) δ 4.95-4.82 (m, 1H), 4.52 (br, 1H), 4.01-3.78 (m, 3H), 3.67 (s, 3H), 3.14-3.00 (m, 2H), 2.96 (s, 3H), 2.09-1.79 (m, 2H), 1.77-1.14 (m, 13H), 1.00-0.85 (m, 6H)

^{13}C NMR (75 MHz, $CDCl_3$) δ 170.6, 165.6, 164.1, 155.9, 79.2, 66.8, 55.9, 52.6, 47.5, 40.1, 38.7, 33.5, 29.4, 28.4, 27.7, 26.3, 23.6, 14.7, 11.9


8
7
6
5
4
3
2
1
ppm


200
180
160
140
120
100
80
60
40
20
0
ppm



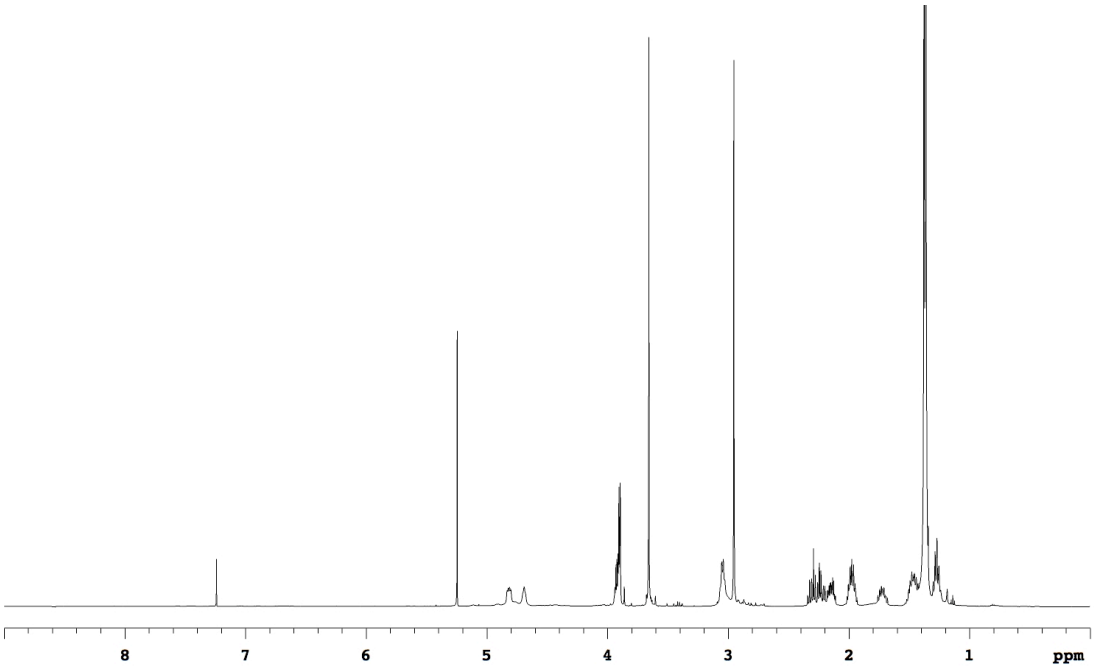
2d

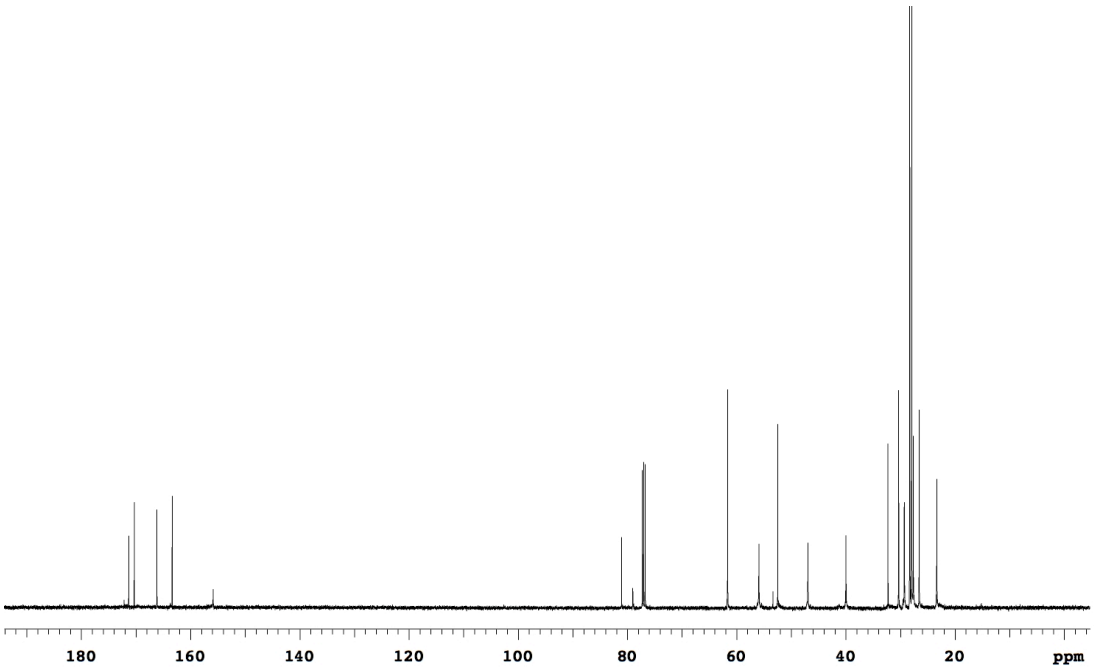
Compound **2d (**1.18 g, 81 %) was prepared from (1.45 g, 2.5 mmol) of Glu-Lys bromide derivative as a white solid.

^{1}H NMR (500 MHz, $CDCl_3$) δ 4.87-4.75 (m, 1H), 4.69 (br, 1H), 3.97-3.87 (m, 3H), 3.66 (s, 3H), 3.13-2.98 (m, 2H), 2.95 (s, 3H), 2.36-2.10 (m, 3H), 2.04-1.92 (m, 2H), 1.78-1.67 (m, 1H), 1.54-1.32 (m, 20H), 1.31-1.22 (m, 2H)

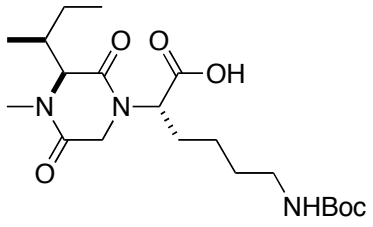
^{13}C NMR (125 MHz, $CDCl_3$) δ 171.3, 170.3, 166.2, 163.4, 155.9, 81.1, 79.0, 61.7, 55.9, 52.5, 46.9, 40.0, 32.3, 30.3, 29.3, 28.3, 27.9, 27.6, 26.7, 23.3

MS (ESI, m/z) 500.3 $(M+H)^+$





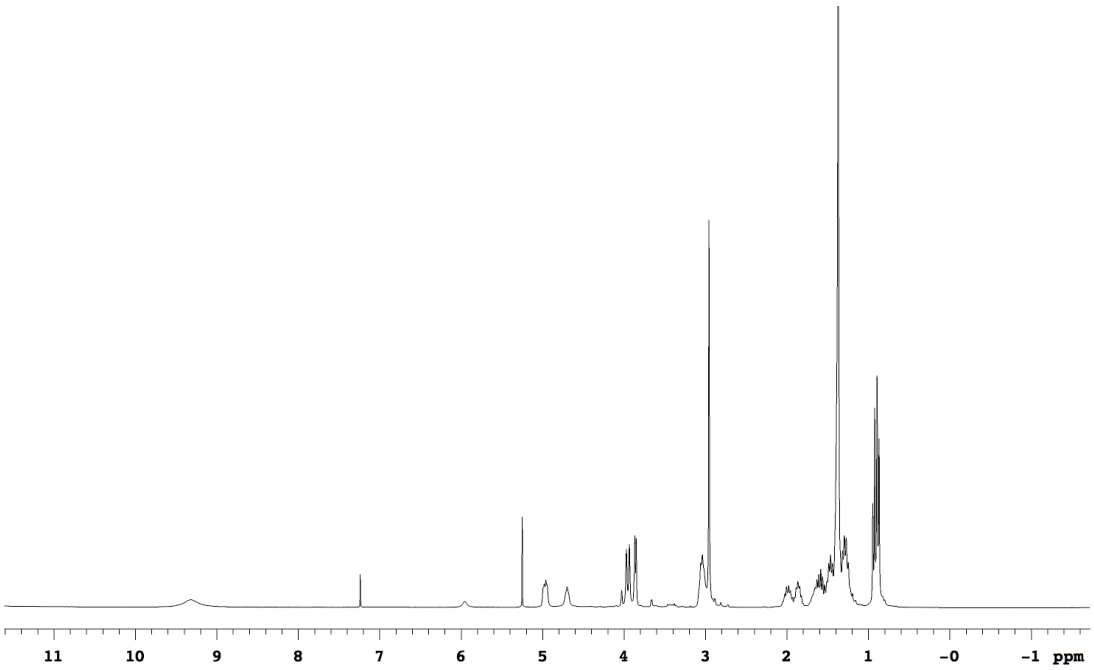
General Procedure for Preparation of 3c-d. The carboxylic ester (1 equiv) was dissolved in dry 1,2-dichloroethane (0.10 M) and after addition of trimethyltin hydroxide (3.0 equiv), the mixture was heated at 80 °C and stirred for 4-6 h under N_2. After completion of the reaction, the mixture was concentrated *in vacuo*, and the residue wad taken up in EtOAc (20 ml). The organic layer was washed with 5 % HCl (3×30 ml) and brine (30 ml), and dried over Na_2SO_4, and concentrated to dryness. Flash chromatography with MeOH/CH_2Cl_2 mixtures provided pure carboxylic acid **3c-d**.

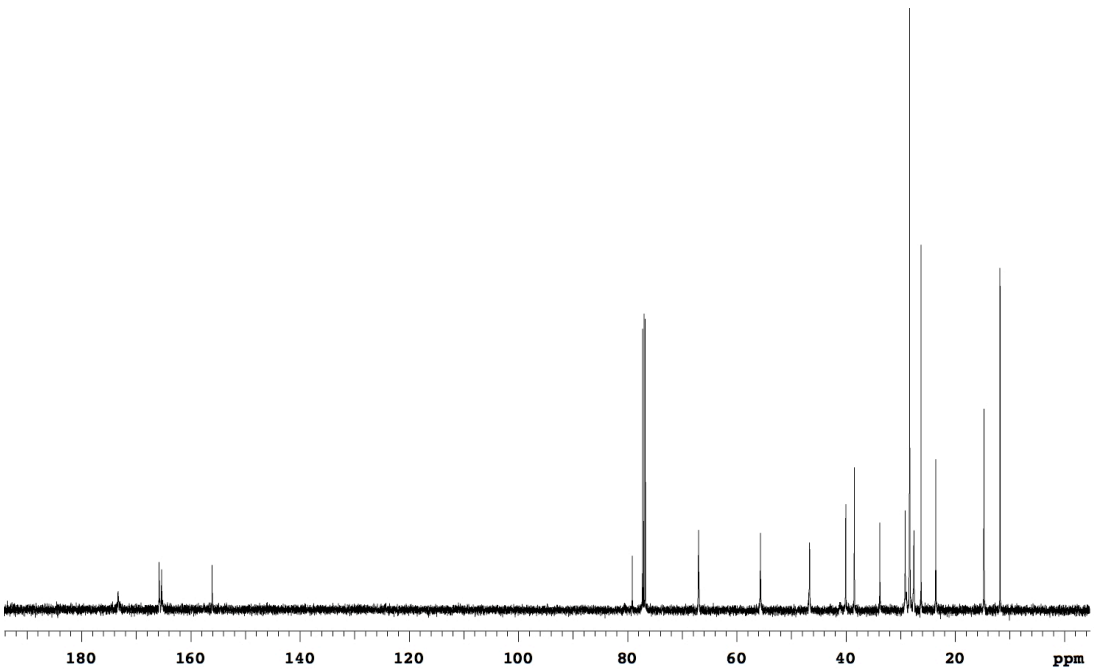


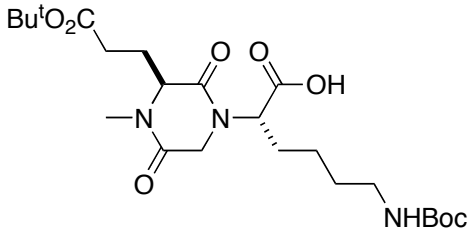
3c

Compound **3c** was prepared from (0.53 g, 1.2 mmol) of Ile-Lys methyl ester **2c**. Flash chromatography (2 % to 10 % MeOH/CH_2Cl_2) afforded 0.34 g (66 %) **3c** as a white solid.

^{1}H NMR (300 MHz, $CDCl_3$) δ 9.31 (br, 1H), 5.06-4.88 (m, 1H), 4.70 (br, 1H), 4.07-3.78 (m, 3H), 3.14-2.86 (m, 5H), 2.08-1.78 (m, 2H), 1.74-1.13 (m, 13H), 1.01-0.76 (m, 6H)
^{13}C NMR (125 MHz, $CDCl_3$) δ 173.3, 165.8, 165.3, 156.1, 79.1, 67.0, 55.7, 46.7, 40.0, 38.5, 33.8, 29.2, 28.3, 27.6, 26.2, 23.6, 14.7, 11.8







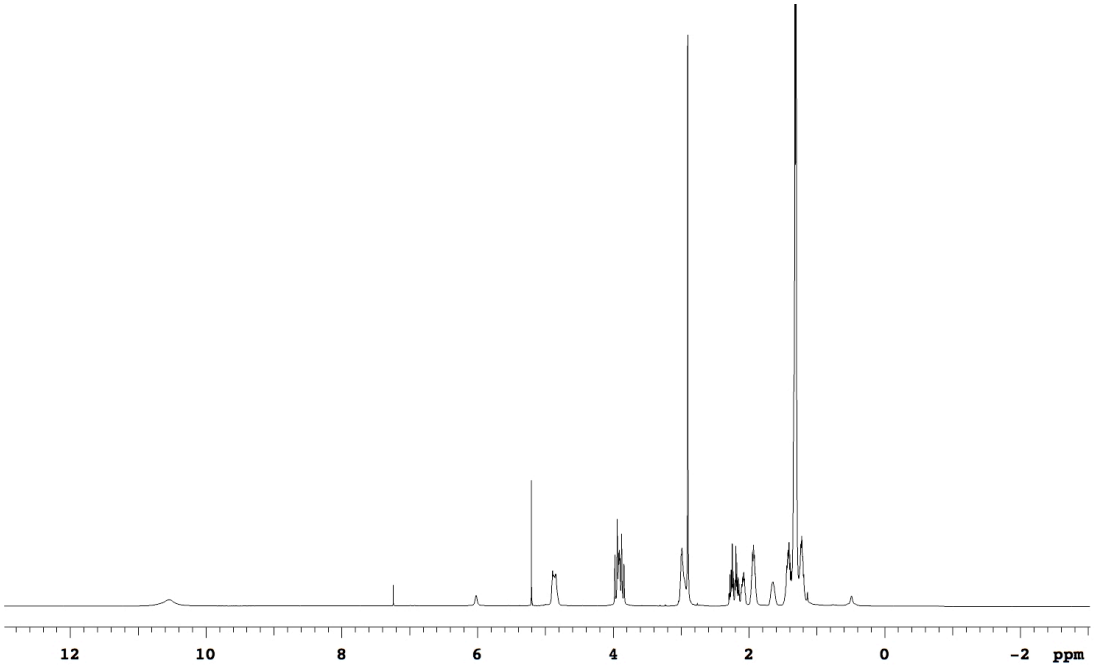
3d

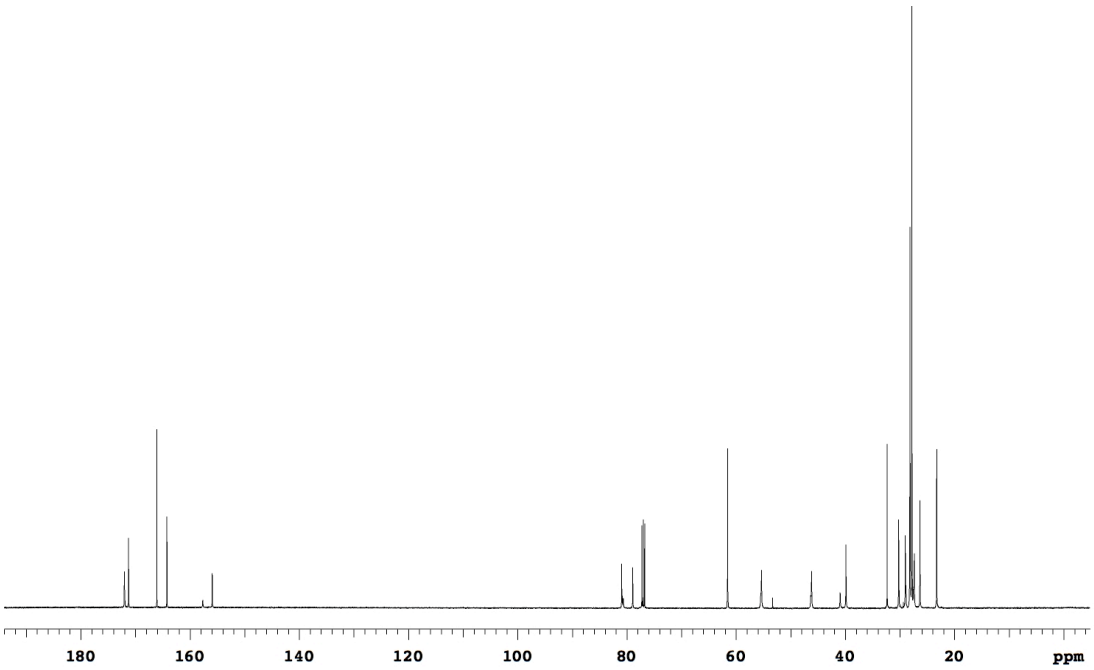
Compound **3d** was prepared from (0.82 g, 1.6 mmol) of Glu-Lys methyl ester **2d**. Flash chromatography (3 % to 5 % MeOH/CH_2Cl_2) afforded 0.50 g (64 %) **3d** as a white solid.

^{1}H NMR (500 MHz, $CDCl_3$) δ 10.54 (br, 1H), 4.95-4.78 (m, 2H), 4.01-3.79 (m, 3H), 3.07-2.81 (m, 5H), 2.35-2.02 (m, 3H), 2.00-1.85 (m, 2H), 1.71-1.57 (m, 1H), 1.51-1.16 (m, 22H)

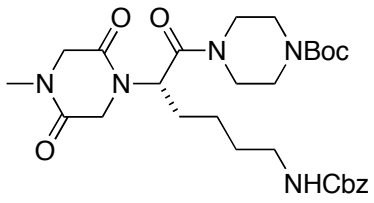
^{13}C NMR (125 MHz, $CDCl_3$) δ 172.0, 171.3, 166.1, 164.2, 155.9, 81.0, 78.9, 61.6, 55.4, 46.2, 39.9, 32.3, 30.2, 29.0, 28.2, 27.8, 27.4, 26.3, 23.3

MS (ESI, m/z) 484.3 $(M-H)^-$





General Procedure for Preparation of 4a-d and 5a-d. The diketopiperazine carboxylic acid **3a-d** (1 equiv) and short or long linker (1 equiv) were dissolved in dry CH_2Cl_2 (0.15 M). *N*-methylmorpholine (2.1 equiv) was added to the above solution followed with HOBt (1.1 equiv) and EDC (1.1 equiv). The resulting solution was stirred at room temperature for 12h. The resulting mixture was washed with H_2O (30 ml), 1N HCl (30 ml), saturated $NaHCO_3$ (30 ml) and brine (30 ml). The organic layer was separated and dried with Na_2SO_4 and concentrated to dryness to afford the crude product. Flash chromatography with MeOH/CH_2Cl_2 mixtures provided pure **4a-d**.



4a

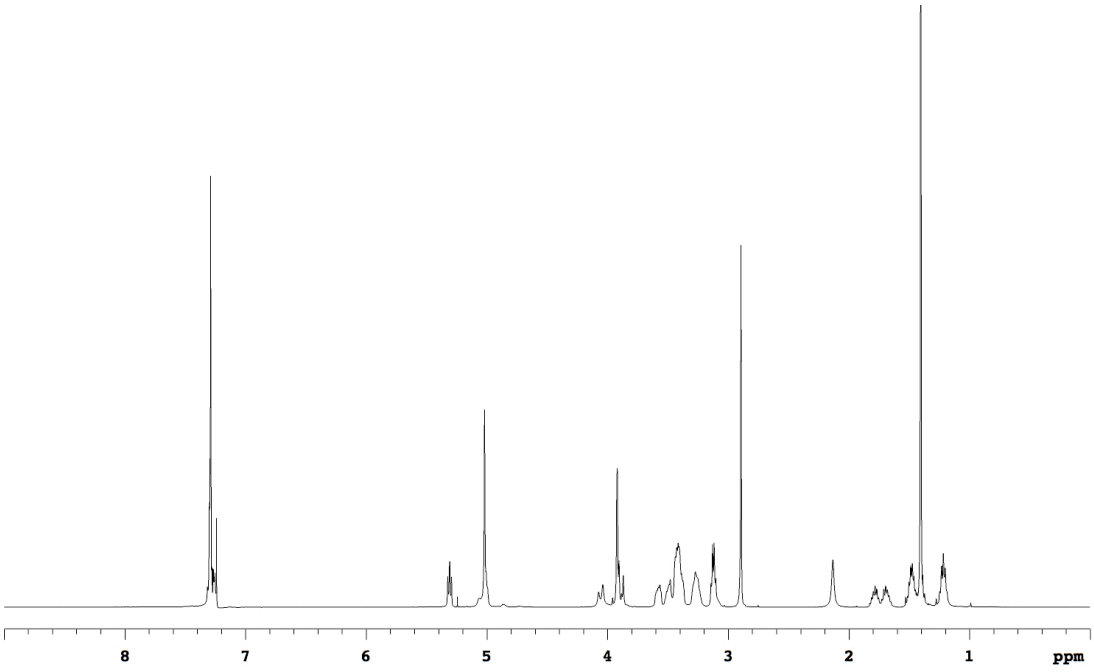
Compound **4a** was prepared from (0.31 g, 0.7 mmol) of cyclo(Gly-Lys) *tert*-butyl ester **2a** and 1-Boc-piperazine. **3a** (1 equiv) was first dissolve in a mixture of dry CH_2Cl_2 (32 equiv), TFA (13 equiv) and triisopropyl silane (2.5 equiv). The resulting mixture was

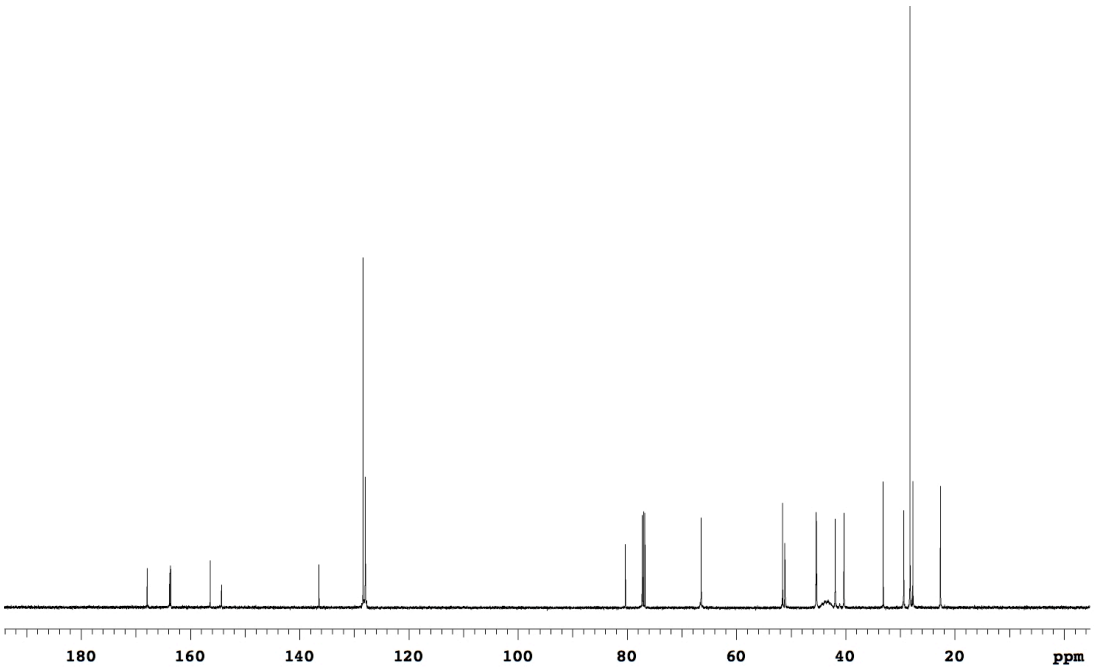
stirred at room temperature for 10 h. The solvents were removed under reduced pressure. The crude product was used for the next reaction following the above general procedure for **4** and **5** without further purification. Flash chromatography (2 % to 5 % $MeOH/CH_2Cl_2$) afforded 0.26 g (66 %) **4a** as a white solid.

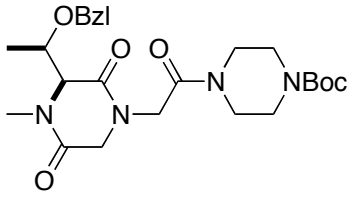
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.31-7.23 (m, 5H), 5.31 (t, 1H, J = 7.7 Hz), 5.10-4.96 (m, 3H), 4.06 (d, 1H, J = 17.4 Hz), 3.95-3.85 (m, 3H), 3.64-3.54 (m, 1H), 3.54-3.46 (m, 1H), 3.46-3.34 (m, 4H), 3.33-3.19 (m, 2H), 3.18-3.06 (m, 2H), 2.89 (s, 3H), 1.94-1.74 (m, 1H), 1.74-1.62 (m, 1H), 1.55-1.35 (m, 11H), 1.27-1.14 (m, 2H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 167.9, 163.8, 163.6, 156.4, 154.3, 136.4, 128.4, 128.0, 128.0, 80.3, 66.4, 51.5, 51.1, 45.4, 45.3, 41.9, 40.3, 33.1, 29.4, 28.2, 27.7, 22.6

MS (APCI, m/z) 559.9 $(M+H)^+$







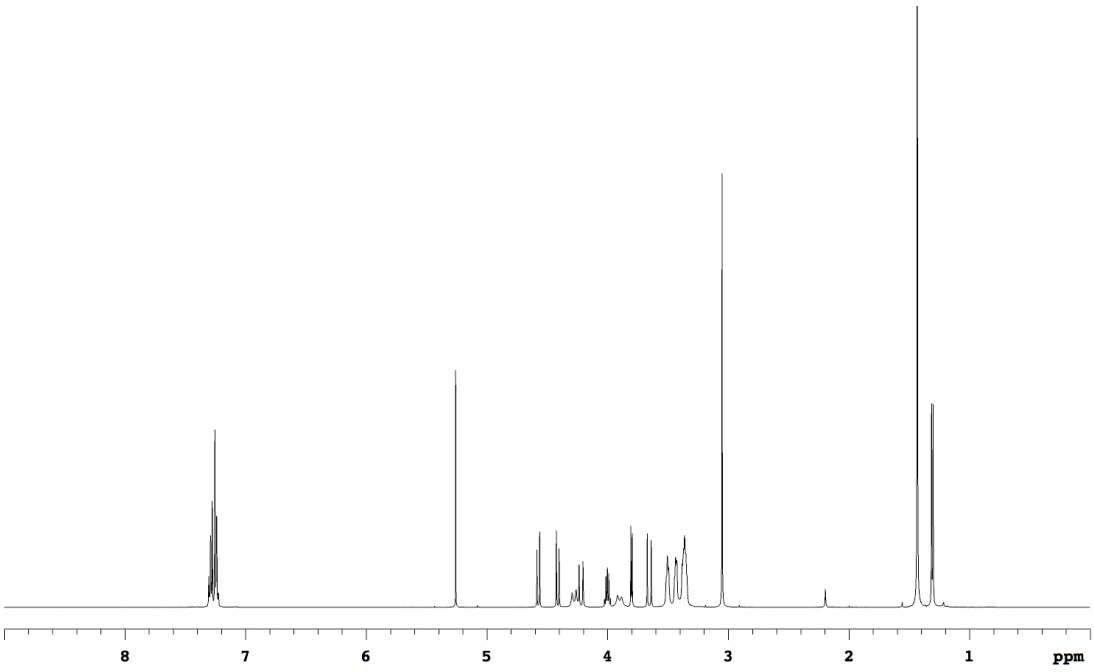
4b

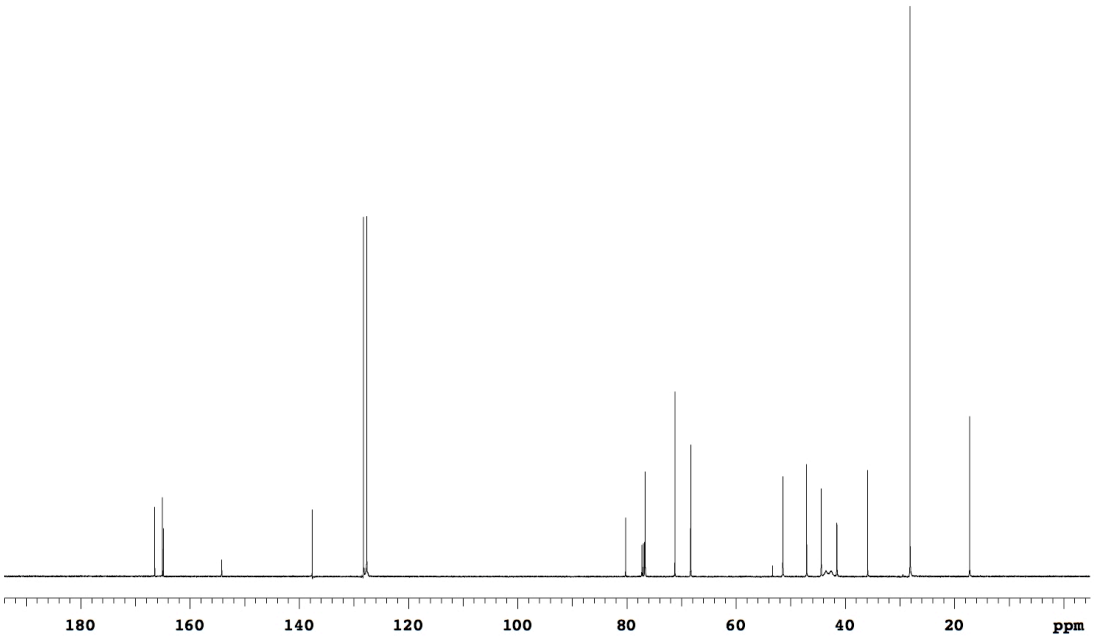
Compound **4b** was prepared from (0.19 g, 0.5 mmol) of cyclo(Thr-Gly) *tert*-butyl ester **2b** and 1-Boc-piperazine following the same procedure with **4a**. Flash chromatography (2 % to 5 % MeOH/CH_2Cl_2) afforded 0.16 g (65 %) **4b** as a colorless oil.

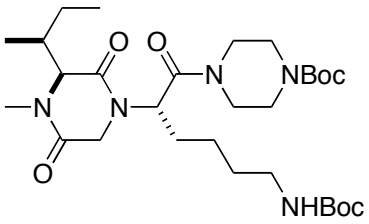
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.33-7.21 (m, 5H), 4.57 (d, 1H, J = 11.6 Hz), 4.41 (d, 1H, J = 11.6 Hz), 4.28 (d, 1H, J = 16.0 Hz), 4.22 (d, 1H, J = 16.5 Hz), 4.04-3.97 (m, 1H), 3.90 (d, 1H, J = 16.0 Hz), 3.80 (d, 1H, J = 5.4 Hz), 3.65 (d, 1H, J = 16.5 Hz), 3.54-3.47 (m, 2H), 3.46-3.41 (m, 2H), 3.40-3.32 (m, 4H), 3.05 (s, 3H), 1.43 (s, 9H), 1.31 (d, 3H, J = 6.4 Hz)

^{13}C NMR (125 MHz, $CDCl_3$) δ 166.5, 165.0, 164.9, 154.2, 137.6, 128.3, 127.6, 127.6, 80.2, 76.6, 71.2, 68.3, 51.4, 47.1, 44.4, 41.5, 35.9, 28.1, 17.2

MS (ESI, m/z) 489.3 $(M+H)^+$


8
7
6
5
4
3
2
1
ppm


180
160
140
120
100
80
60
40
20
ppm



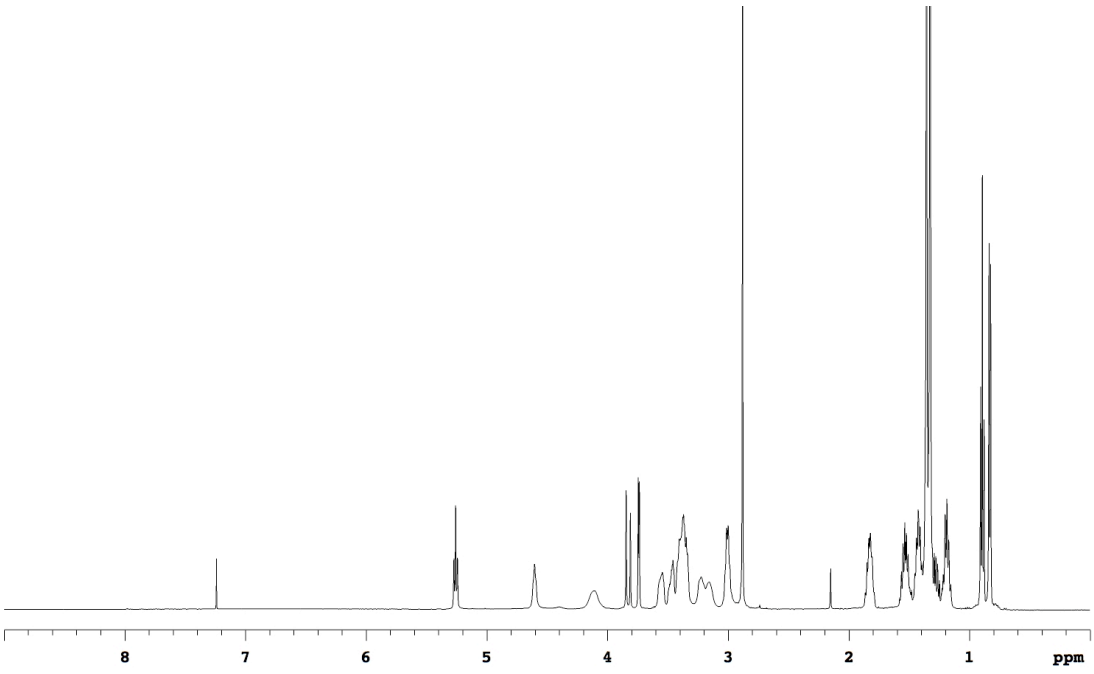
4c

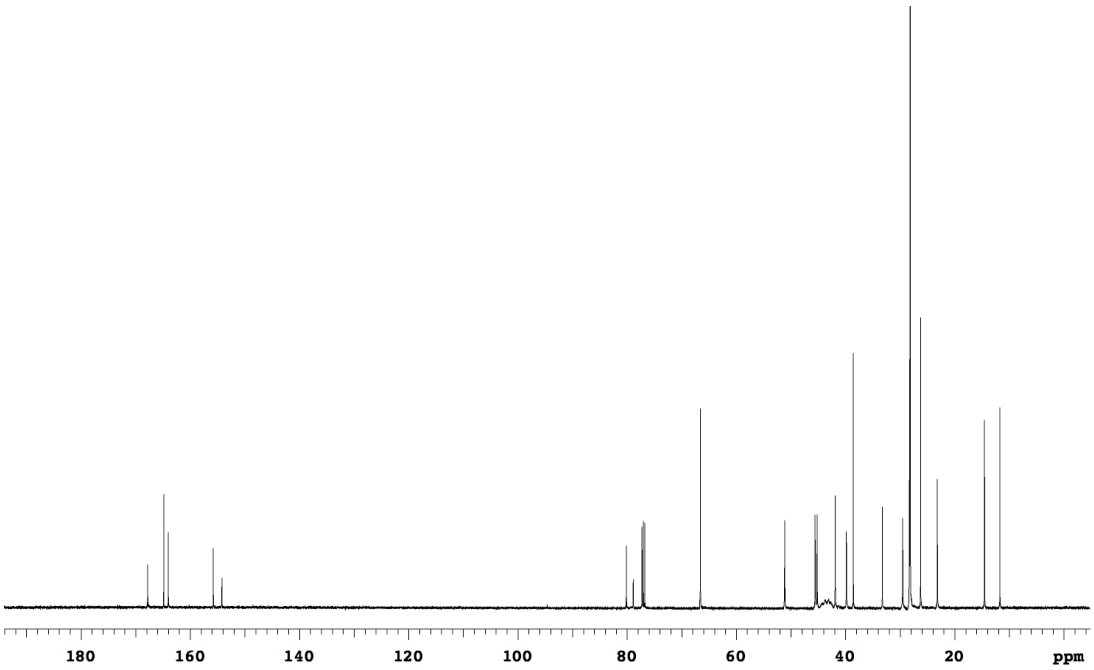
Compound **4c** was prepared from (0.26 g, 0.6 mmol) of cyclo(Ile-Lys) carboxylic acid **3c** and 1-Boc-piperazine. Flash chromatography (2 % to 3 % MeOH/CH_2Cl_2) afforded 0.30 g (86 %) **4c** as a white solid.

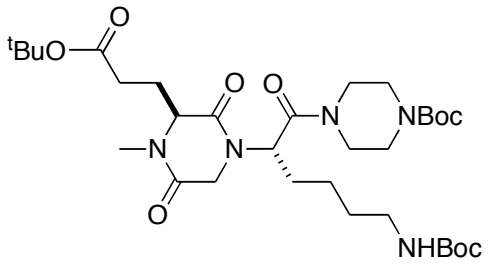
^{1}H NMR (500 MHz, $CDCl_3$) δ 5.26 (t, 1H, J = 7.6 Hz), 4.60 (br, 1H), 4.11 (br, 1H), 3.83 (d, 1H, J = 17.6 Hz), 3.74 (d, 1H, J = 4.3 Hz), 3.60-3.31 (m, 6H), 3.27-3.11 (m, 2H), 3.04-2.97 (m, 2H), 2.88 (s, 3H), 1.88-1.78 (m, 2H), 1.59-1.47 (m, 2H), 1.47-1.39 (m, 2H), 1.35 (s, 9H), 1.33 (s, 9H), 1.31-1.24 (m, 1H), 1.24-1.14 (m, 2H), 0.89 (t, 3H, J = 7.3 Hz), 0.83 (d, 3H, J = 6.8 Hz)

^{13}C NMR (125 MHz, $CDCl_3$) δ 167.8, 162.9, 164.0, 155.8, 154.2, 80.1, 78.9, 66.6, 51.1, 45.6, 45.2, 41.9, 39.8, 38.6, 33.3, 29.5, 28.3, 28.2, 28.1, 26.2, 23.2, 14.5, 11.7

MS (ESI, m/z) 582.4 $(M+H)^+$







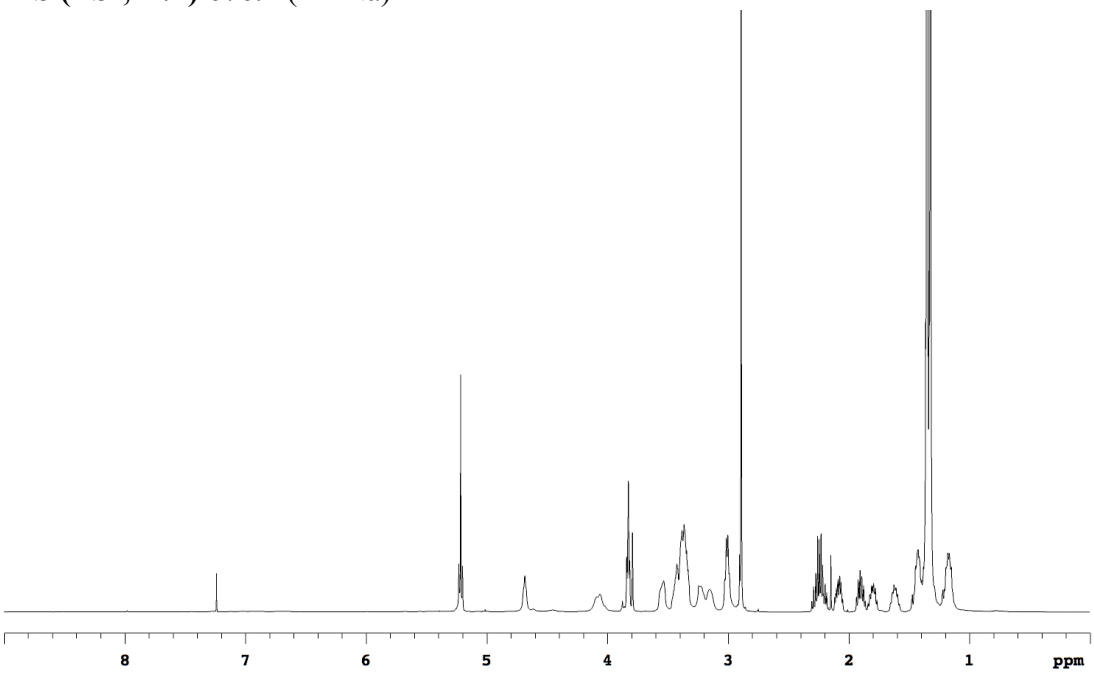
4d

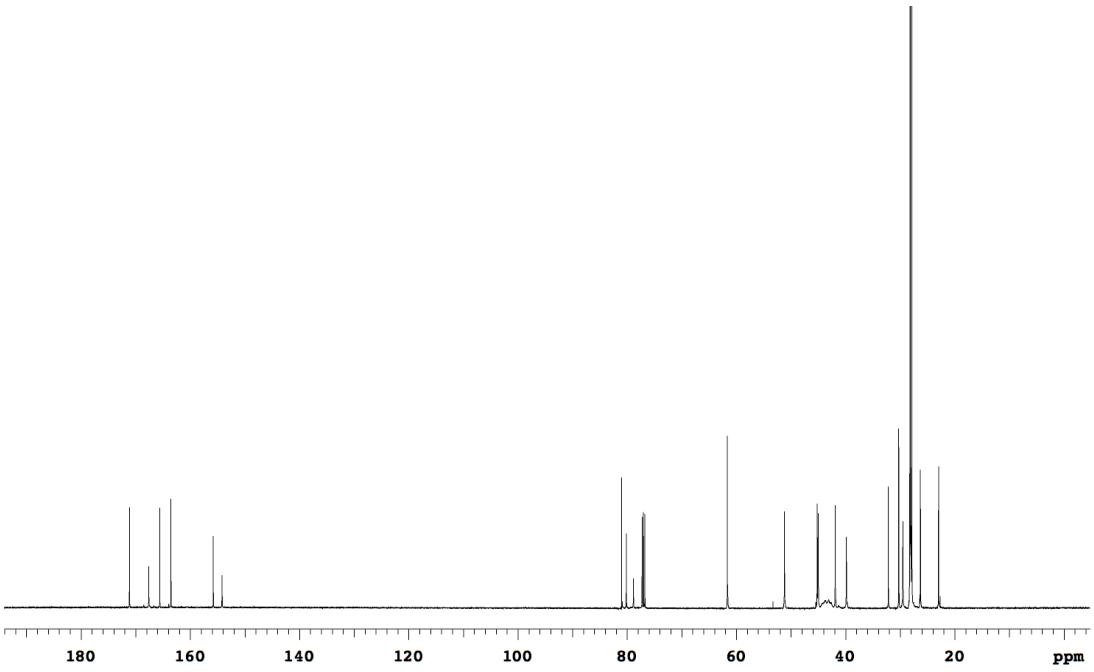
Compound **4d** was prepared from (0.22 g, 0.4 mmol) of cyclo(Glu-Lys) carboxylic acid **3d** and 1-Boc-piperazine. Flash chromatography (2 % to 3 % MeOH/CH_2Cl_2) afforded 0.22 g (75 %) **4b** as a white solid.

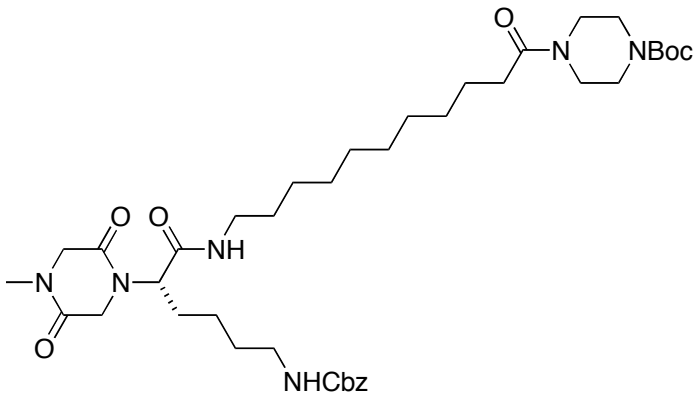
^{1}H NMR (500 MHz, $CDCl_3$) δ 5.22 (t, 1H, J = 7.6 Hz), 4.69 (br, 1H), 4.10(br, 1H), 3.88-3.78(m, 2H), 3.60-3.09 (m, 8H), 3.05-2.96 (m, 2H), 2.89 (s, 3H), 2.30-2.17 (m, 2H), 2.13-2.04 (m, 1H), 1.95-1.85 (m, 1H), 1.85-1.76 (m, 1H), 1.67-1.57 (m, 1H), 1.49-1.28 (m, 29H), 1.23-1.12 (m, 2H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 171.1, 167.6, 165.6, 163.5, 155.8, 154.2, 81.0, 80.1, 78.8, 61.6, 51.2, 45.2, 45.0, 41.9, 39.8, 32.2, 30.3, 29.5, 28.2, 28.1, 28.0, 27.9, 26.3, 23.0

MS (ESI, m/z) 676.4 $(M+Na)^+$







5a

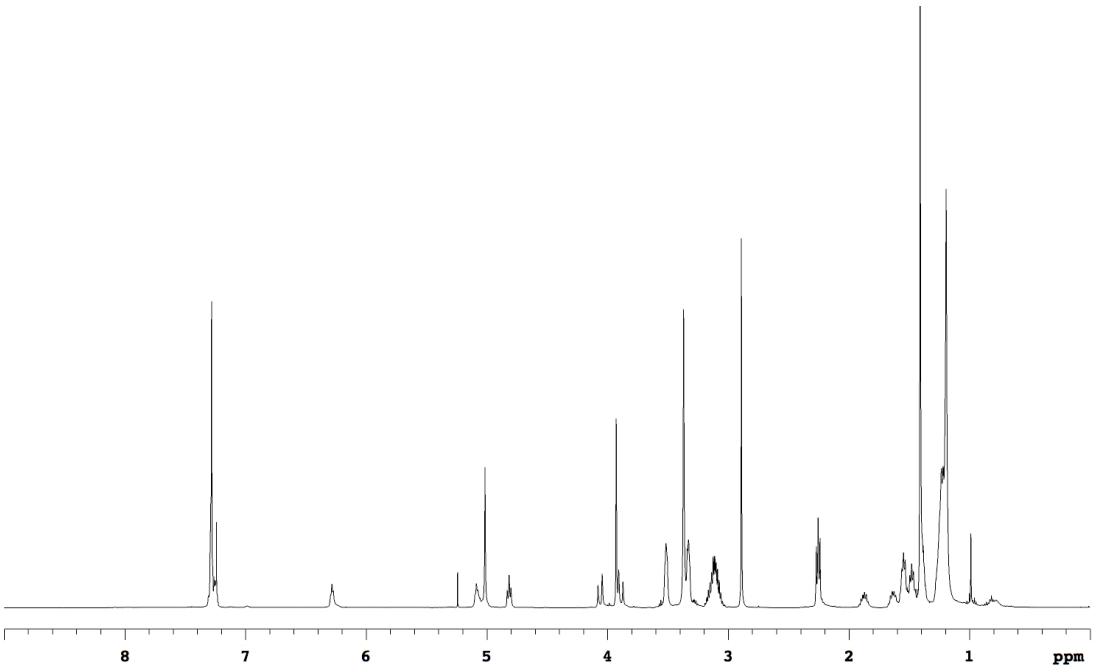
Compound **5a** was prepared from (0.27 g, 0.6 mmol) of cyclo(Gly-Lys) *tert*-butyl ester **2b** and long linker following the same procedure with **4a**.

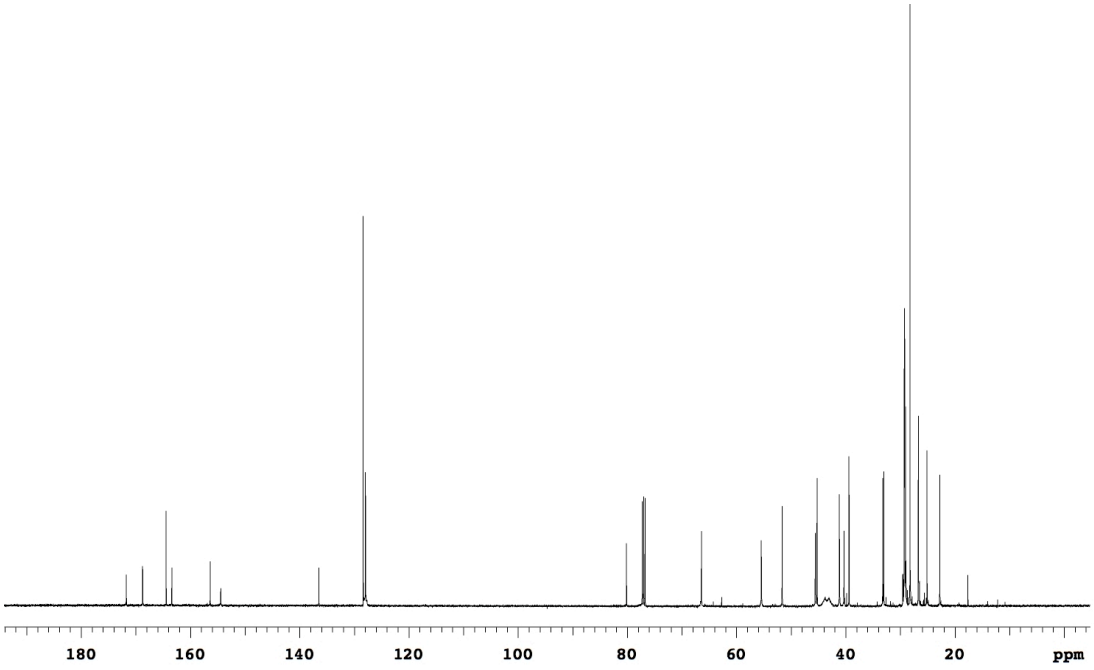
Flash chromatography (2 % to 5 % MeOH/CH_2Cl_2) afforded 0.27 g (61 %) **5a** as a white solid.

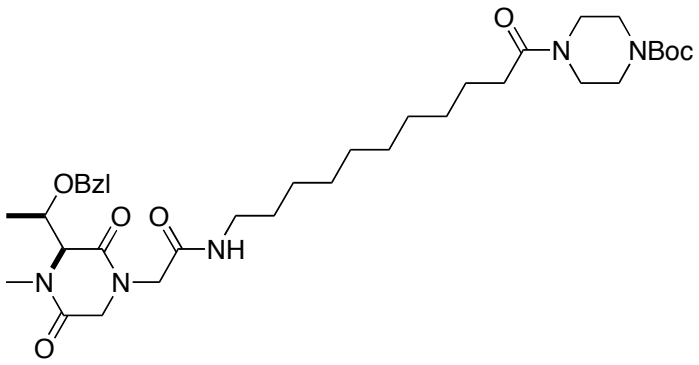
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.33-7.21 (m, 5H), 6.28 (t, 1H, J = 5.4 Hz), 5.14-4.97 (m, 3H), 4.82 (t, 1H, J = 7.6 Hz), 4.06 (d, 1H, J = 17.3 Hz), 3.93 (s, 2H), 3.89 (d, 1H, J = 17.3 Hz), 3.52 (br, 2H), 3.42-3.30 (m, 8H), 3.20-3.02 (m, 4H), 2.89 (s, 3H), 2.25 (t, 2H, J = 7.4 Hz), 1.93-1.81 (m, 1H), 1.69-1.59 (m, 1H), 1.59-1.44 (m, 4H), 1.44-1.35 (m, 11H), 1.31-1.14 (m, 12H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 171.7, 168.8, 164.4, 163.4, 156.4, 154.4, 136.5, 128.4, 128.0, 127.9, 80.1, 66.4, 55.5, 51.6, 45.6, 45.3, 41.2, 40.3, 39.4, 33.2, 33.0, 29.3, 29.3, 29.2, 29.2, 29.2, 29.0, 28.2, 26.7, 26.7, 25.1, 22.8, 17.6

MS (APCI, m/z) 743.2 $(M+H)^+$


8
7
6
5
4
3
2
1
ppm


180
160
140
120
100
80
60
40
20
ppm



5b

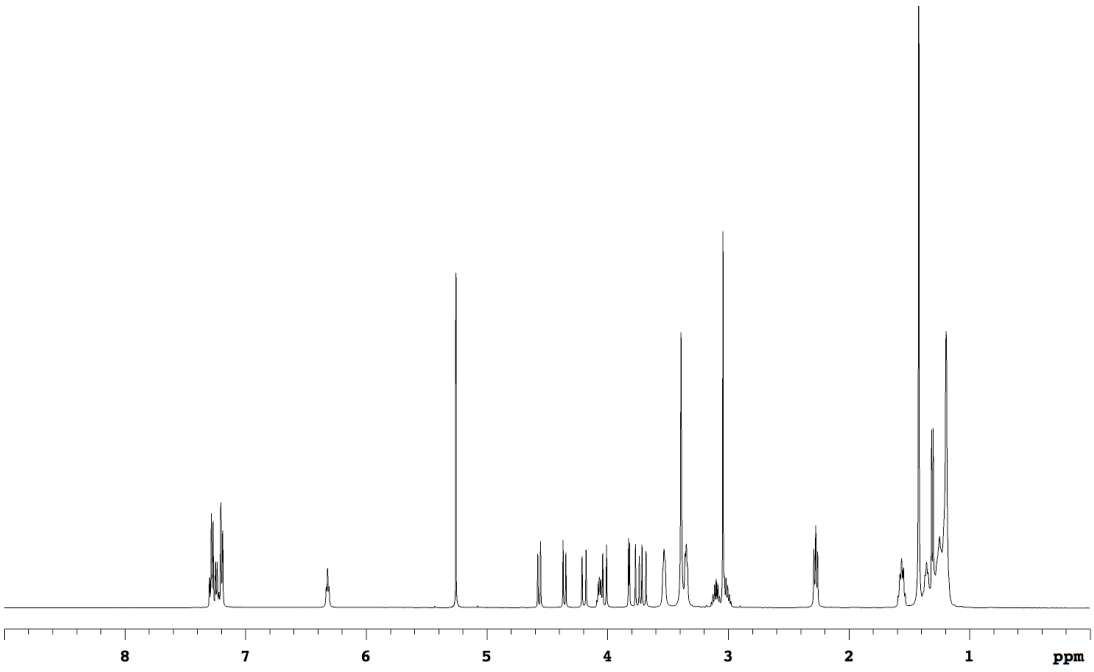
Compound **5b** was prepared from (0.15 g, 0.4 mmol) of cyclo(Thr-Gly) *tert*-butyl ester **2b** and long linker following the same procedure with **4a**.

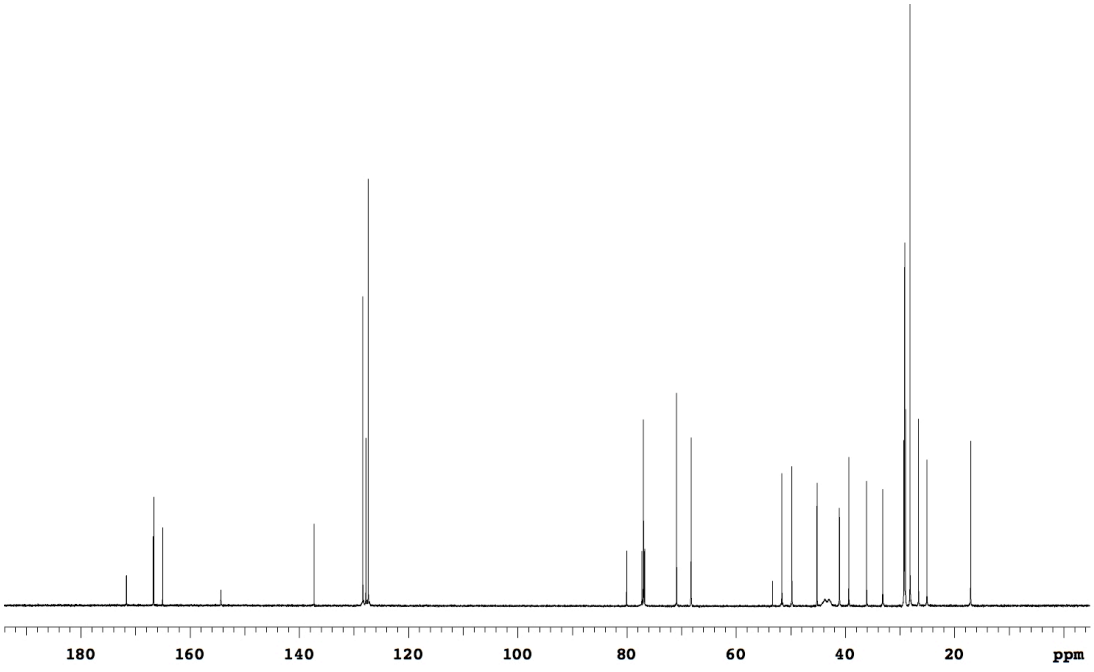
Flash chromatography (2 % to 5 % MeOH/CH_2Cl_2) afforded 0.16 g (60 %) **4b** as a colorless oil.

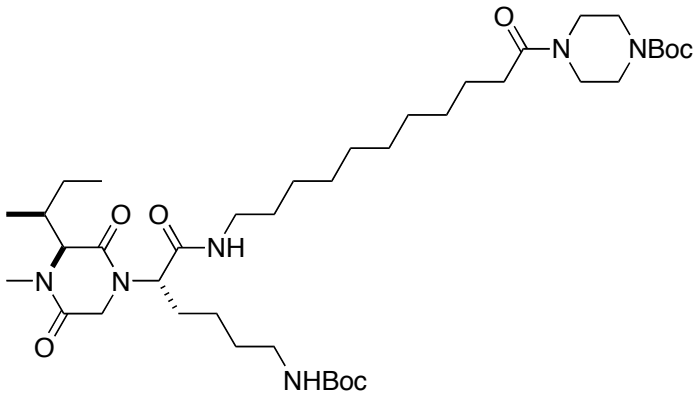
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.33-7.15 (m, 5H), 6.32 (t, 1H, J = 5.6 Hz), 4.57 (d, 1H, J = 11.6 Hz), 4.36 (d, 1H, J = 11.6 Hz), 4.10 (d, 1H, J = 16.6 Hz), 4.10-4.05 (m, 1H), 4.02 (d, 1H, J = 15.9 Hz), 3.82 (d, 1H, J = 4.1 Hz), 3.75 (d, 1H, J = 15.9 Hz), 3.70 (d, 1H, J = 16.6 Hz), 3.58-3.49 (m, 2H), 3.43-3.29 (m, 6H), 3.15-3.07 (m, 1H), 3.06-2.97 (m, 4H), 2.27 (t, 2H, J = 7.6 Hz), 1.61-1.52 (m, 2H), 1.42 (s, 9H), 1.40-1.33 (m, 2H), 1.31 (d, 3H, J = 6.4 Hz), 1.29-1.14 (m, 12H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 171.6, 166.8, 166.7, 165.2, 154.4, 137.3, 128.4, 127.8, 127.4, 80.0, 77.0, 70.9, 68.2, 51.6, 49.8, 45.2, 41.1, 39.4, 36.1, 33.1, 29.3, 29.2, 29.1, 29.0, 28.2, 26.6, 25.1, 17.0

MS (ESI, m/z) 672.4 $(M+H)^+$


8
7
6
5
4
3
2
1
ppm


180
160
140
120
100
80
60
40
20
ppm



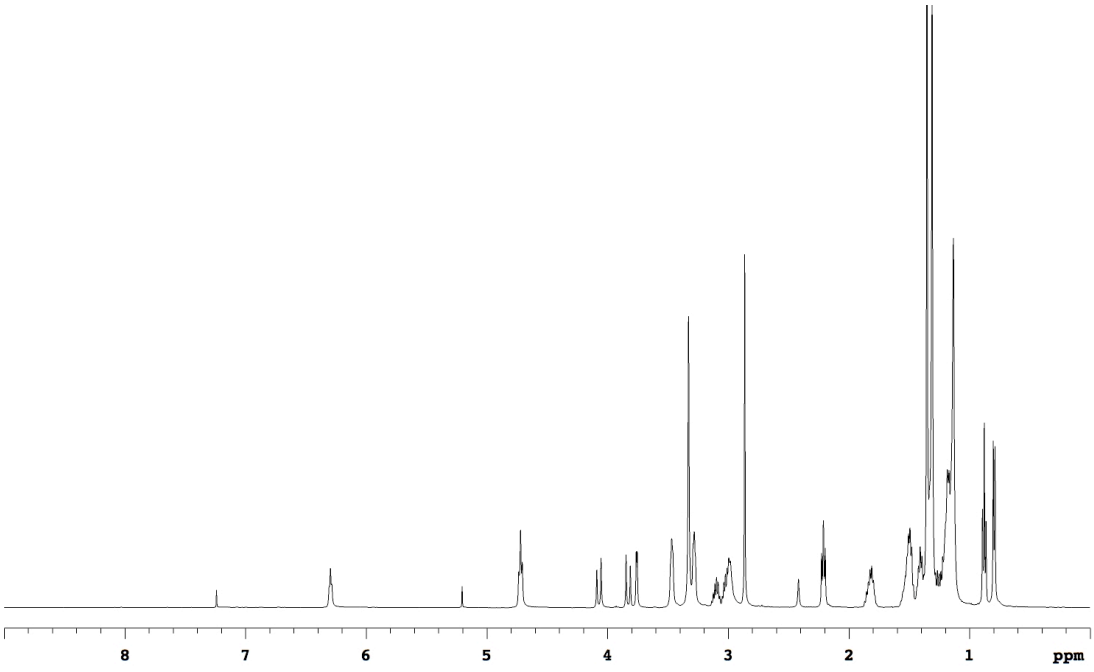
5c

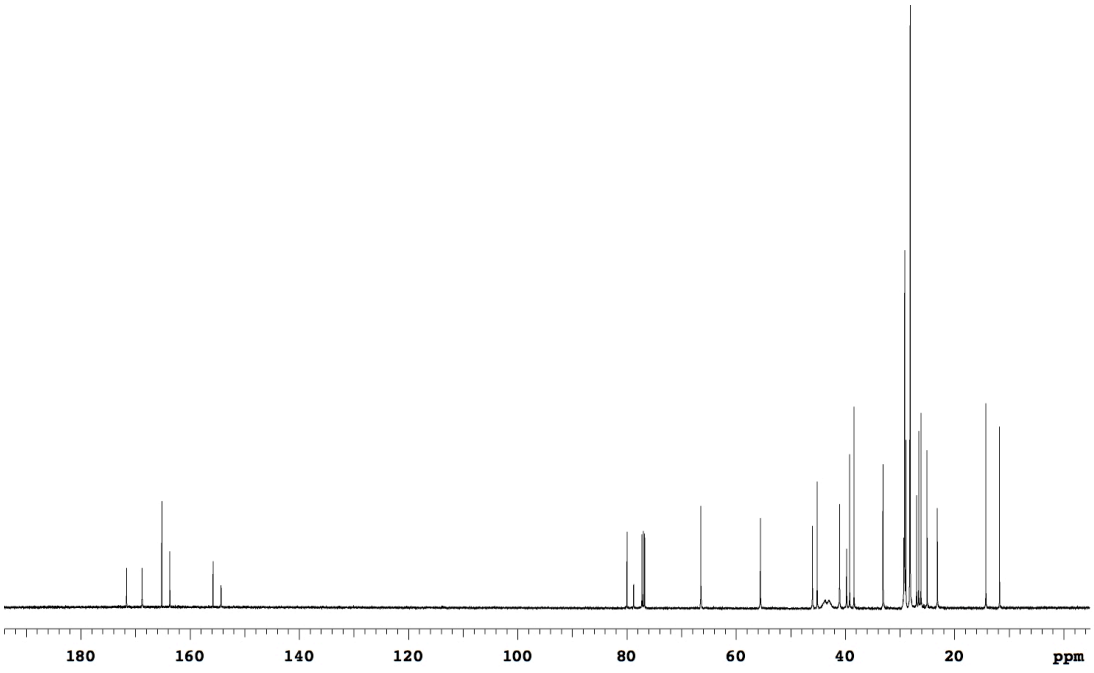
Compound **5c** was prepared from (0.26 g, 0.6 mmol) of cyclo(Ile-Lys) carboxylic acid **3c** and long linker. Flash chromatography (2 % to 3 % MeOH/CH_2Cl_2) afforded 0.38 g (83 %) **5c** as a white solid.

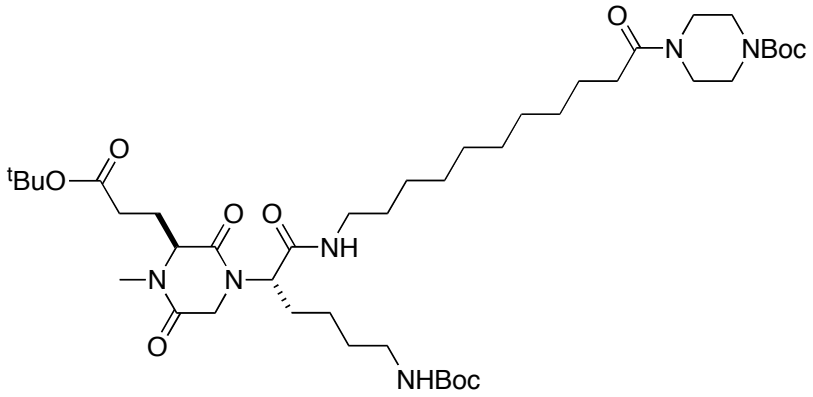
^{1}H NMR (500 MHz, $CDCl_3$) δ 6.30 (t, 1H, J = 5.0 Hz), 4.75-4.69 (m, 2H), 4.07(d, 1H, J = 17.5 Hz), 3.83 (d, 1H, J = 17.5 Hz), 3.76(d, 1H, J = 4.1 Hz), 3.50-3.44 (m, 2H), 3.35-3.25 (m, 6H), 3.15-2.94 (m, 4H), 2.86 (s, 3H), 2.21 (t, 2H, J = 7.5 Hz), 1.88-1.76 (m, 2H), 1.57-1.38 (m, 6H), 1.35 (s, 9H), 1.31(s, 9H), 1.29-1.09 (m, 17H), 0.88(t, 3H, J = 7.2 Hz), 0.80 (d, 3H, J = 6.8 Hz)

^{13}C NMR (125 MHz, $CDCl_3$) δ 171.6, 168.8, 165.2, 163.7, 155.8, 154.3, 80.0, 78.8, 66.4, 55.6, 46.0, 45.1, 41.0, 39.8, 39.2, 38.4, 33.1, 33.1, 29.3, 29.2, 29.1, 29.1, 28.9, 28.2, 28.1, 26.9, 26.5, 26.1, 25.0, 23.1, 14.3, 11.7

MS (ESI, m/z) 765.5 $(M+H)^+$


8
7
6
5
4
3
2
1
ppm


180
160
140
120
100
80
60
40
20
ppm



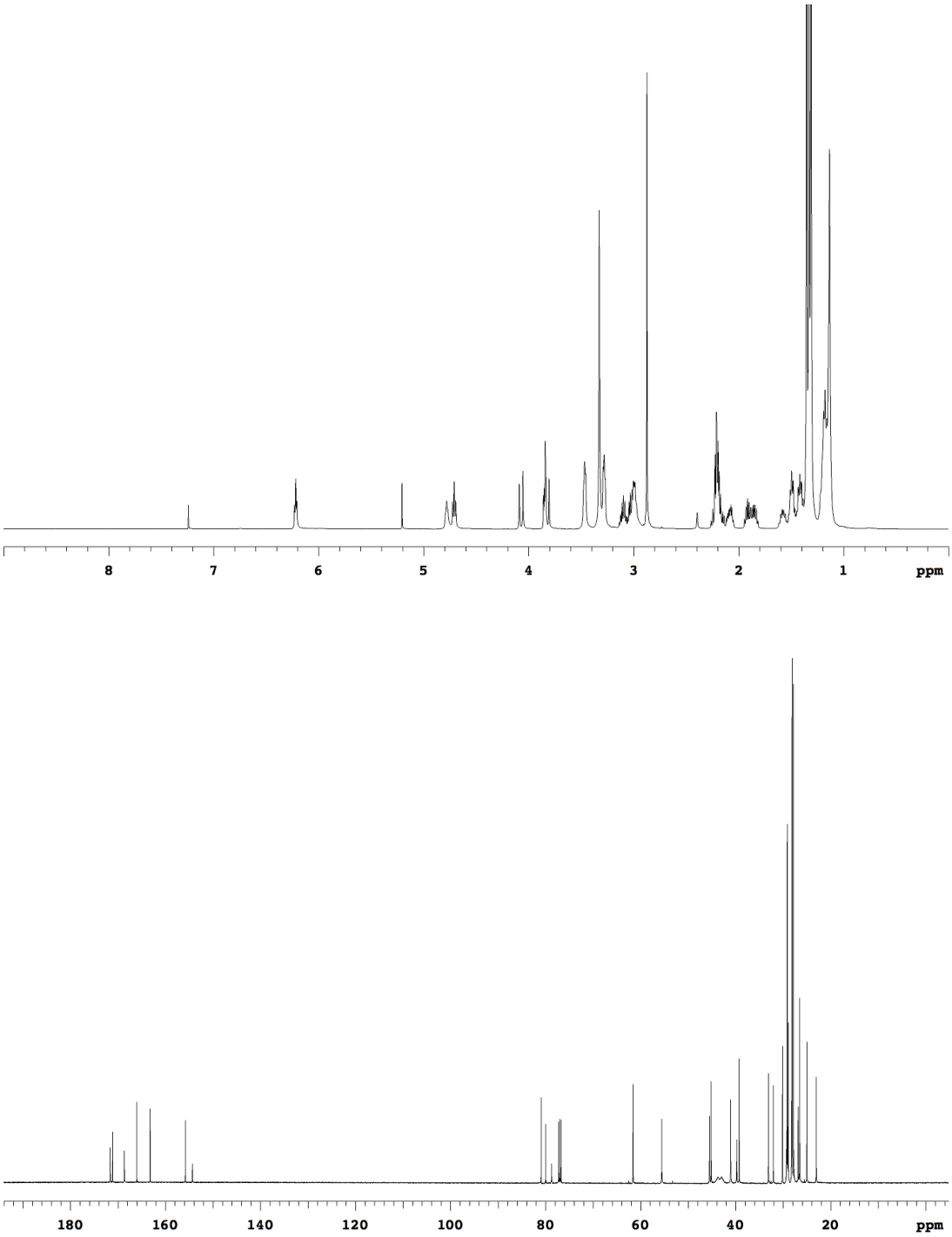
5d

Compound **5d** was prepared from (0.25 g, 0.5 mmol) of cyclo(Glu-Lys) carboxylic acid **3d** and long linker. Flash chromatography (3 % MeOH/CH_2Cl_2) afforded 0.26 g (62 %) **5d** as a white solid.

^{1}H NMR (500 MHz, $CDCl_3$) δ 6.22 (t, 1H, J = 5.6 Hz), 4.82-4.74 (m, 1H), 4.71(t, 1H, J = 7.8 Hz), 4.07 (d, 1H, J = 17.3 Hz), 3.87-3.79 (m, 2H), 3.50-3.44 (m, 2H), 3.35-3.25 (m, 6H), 3.15-2.94 (m, 4H), 2.87 (s, 3H), 2.28-2.04 (m, 5H), 1.96-1.81 (m, 2H), 1.63-1.38 (m, 5H), 1.35 (s, 9H), 1.33 (s, 9H), 1.31 (s, 9H), 1.24-1.09 (m, 16H)

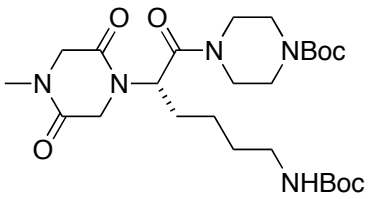
^{13}C NMR (125 MHz, $CDCl_3$) δ 171.6, 171.1, 168.6, 166.0, 163.2, 155.8, 154.3, 80.9, 80.0, 78.7, 61.6, 55.6, 45.5, 45.2, 41.0, 39.8, 39.3, 38.4, 33.1, 32.1, 30.2, 29.3, 29.2, 29.1, 29.1, 28.9, 28.2, 28.1, 27.8, 26.9, 26.5, 26.5, 25.0, 23.0

MS (ESI, m/z) 837.6 $(M+H)^+$



General Procedure for Preparation of 6a and 7a. A solution of diketopiperazine (1 equiv) in dry methanol (0.10 M) was stirred vigorously with 10 % Pd-C (0.1 equiv) in a H_2 atmosphere for 16 h at 25 °C until the reaction is complete.

Filtration and evaporation of the solvent *in vacuo* yielded the crude product and it was used directly for the next reaction. To a solution of di-*tert* butyl dicarbonate (1.0 equiv) in dry CH_2Cl_2 (0.20 M) was added the crude secondary amine from previous step and Et_3N (2.0 equiv). The reaction mixture was stirred at 25 °C for 12 h. The mixture was washed with H_2O (30 ml), 1N HCl (30 ml), saturated $NaHCO_3$ (30 ml) and brine (30 ml). The organic layer was separated and dried with Na_2SO_4 and concentrated to dryness to give the crude product. Flash chromatography with MeOH/CH_2Cl_2 mixtures provided pure **6a and 7a**.



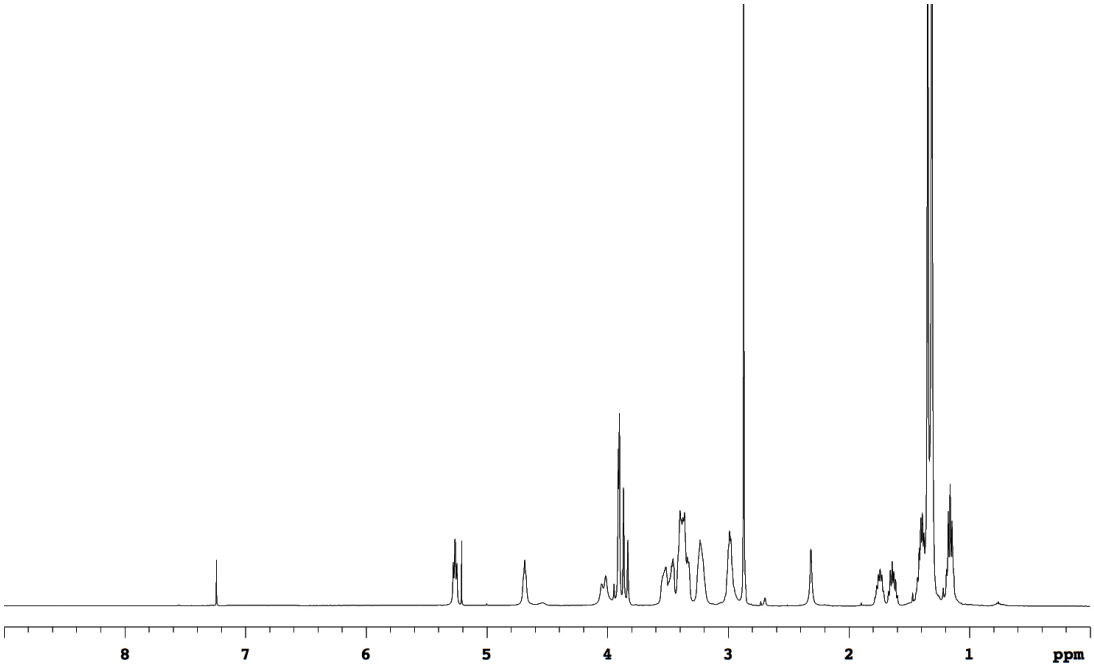
6a

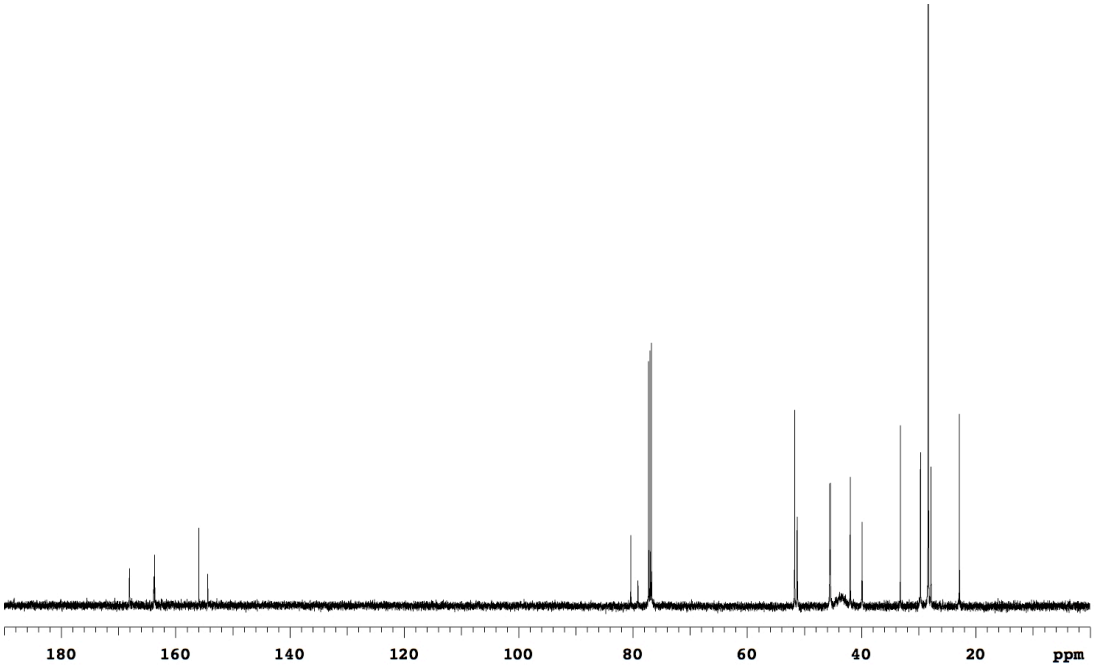
Compound **6a** was prepared from (0.17 g, 0.4 mmol) of cyclo(Gly-Lys) derivative **4a**. Flash chromatography (2 % to 3 % MeOH/CH_2Cl_2) afforded 0.17 g (72 %) **6a** as a white solid.

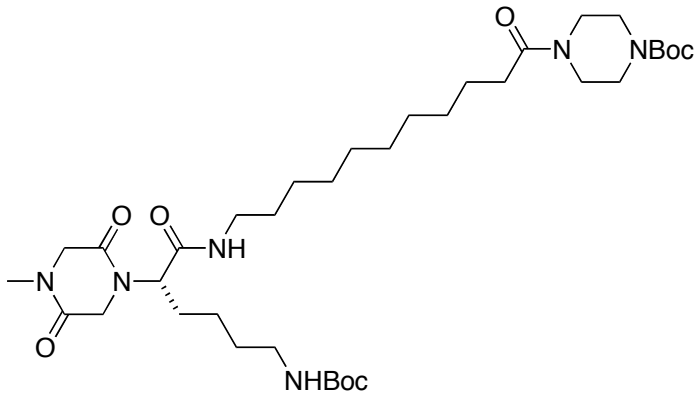
^{1}H NMR (500 MHz, $CDCl_3$) δ 5.26 (t, 1H, J = 7.6 Hz), 4.69 (br, 1H), 4.03 (d, 1H, J = 17.4 Hz), 3.96-3.81 (m, 3H), 3.58-3.30 (m, 6H), 3.27-3.18 (m, 2H), 3.03-2.93 (m, 2H), 2.87 (s, 3H), 1.80-1.58 (m, 2H), 1.45-1.37 (m, 2H), 1.35 (s, 9H), 1.31 (s, 9H), 1.20-1.12 (m, 2H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 168.1, 163.8, 163.7, 155.9, 154.4, 80.4, 79.1, 51.7, 51.3, 45.6, 45.4, 42.0, 39.9, 33.2, 29.7, 28.4, 28.3, 27.8, 22.9

MS (ESI, m/z) 548.3 $(M+Na)^+$


8
7
6
5
4
3
2
1
ppm


180
160
140
120
100
80
60
40
20
ppm



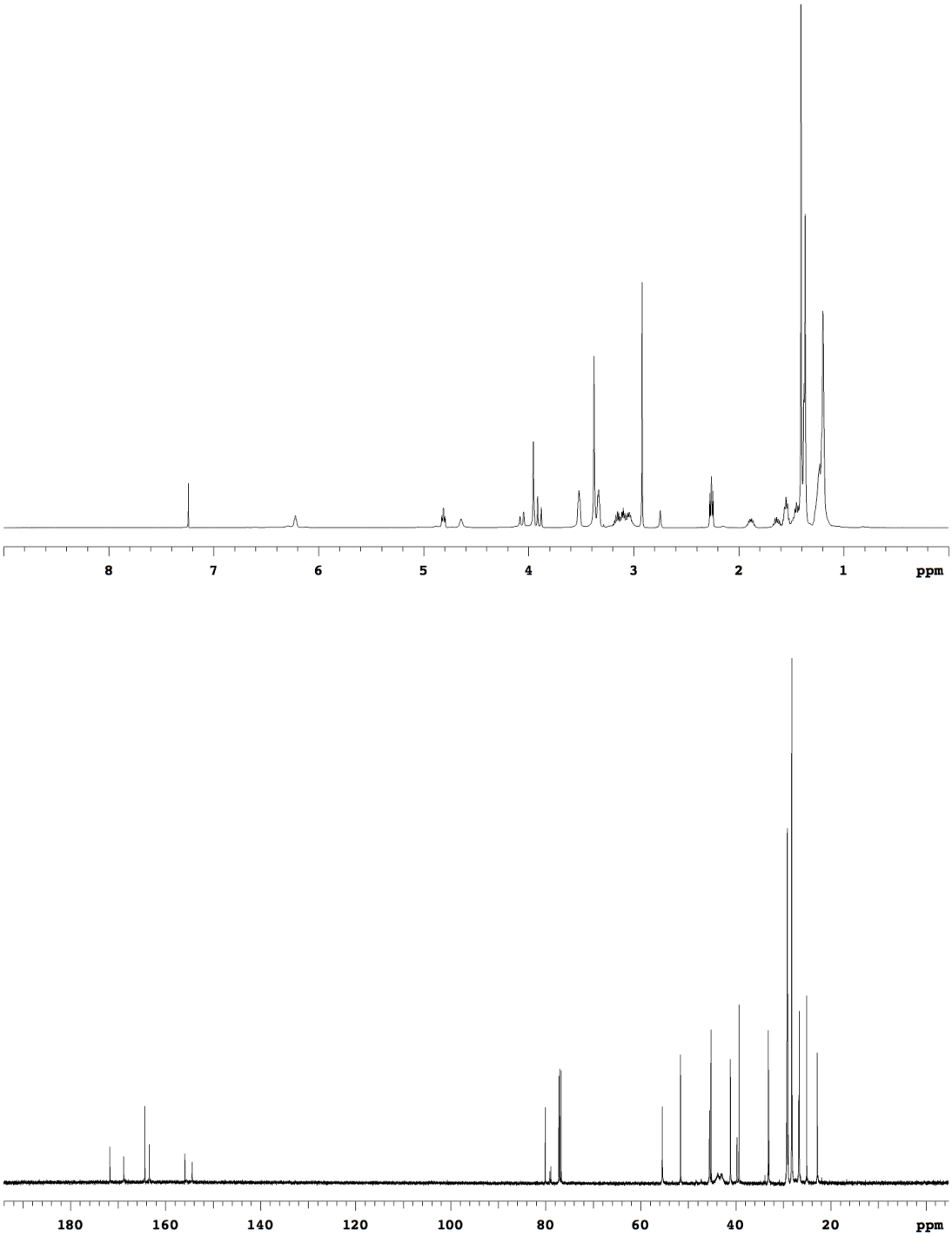
6b

Compound **6b** was prepared from (0.21 g, 0.4 mmol) of cyclo(Gly-Lys) derivative **5b**. Flash chromatography (2 % to 10 % MeOH/CH_2Cl_2) afforded 0.20 g (81 %) **6b** as a white solid.

^{1}H NMR (500 MHz, $CDCl_3$) δ 6.22 (br, 1H), 4.81 (t, 1H, J = 7.7 Hz), 4.64 (br, 1H), 4.06 (d, 1H, J = 17.7 Hz), 3.96 (s, 2H), 3.90 (d, 1H, J = 17.7 Hz), 3.52 (br, 2H), 3.38 (br, 4H), 3.36-3.31 (m, 2H), 3.22-2.98 (m, 4H), 2.92 (s, 3H), 2.26 (t, 2H, J = 7.7 Hz), 1.94-1.82 (m, 1H), 1.69-1.59 (m, 1H), 1.59-1.51 (m, 2H), 1.51-1.32 (m, 22H), 1.30-1.12 (m, 14H)

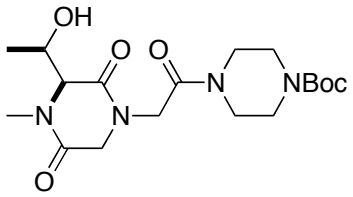
^{13}C NMR (125 MHz, $CDCl_3$) δ 171.7, 168.8, 164.4, 163.4, 155.9, 154.4, 80.1, 79.1, 78.9, 55.4, 51.6, 45.6, 45.2, 41.1, 39.3, 33.2, 33.0, 29.3, 29.2, 29.2, 29.2, 29.2, 29.0, 28.3, 28.2, 28.2, 26.7, 26.6, 25.1, 22.8

MS (ESI, m/z) 709.5 $(M+H)^+$



General Procedure for Preparation of 6b and 7b. A solution of Bzl protected diketopiperazine derivative (1 equiv) in dry methanol (0.10 M) was stirred vigorously

with 10 % Pd-C (0.1 equiv) in a H_2 atmosphere for 12 h at 25 °C until the reaction is complete. Filtration and evaporation of the solvent *in vacuo* yielded the crude product. The residue was purified by flash chromatography with MeOH/CH_2Cl_2 mixtures to provid pure **6b and 7b**.



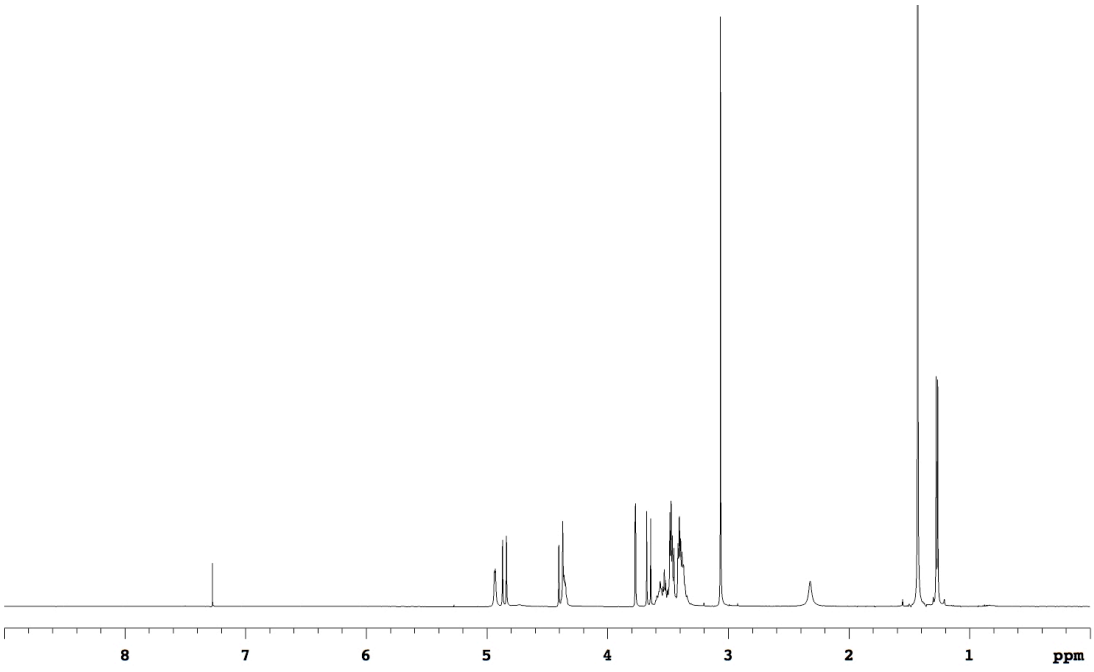
6b

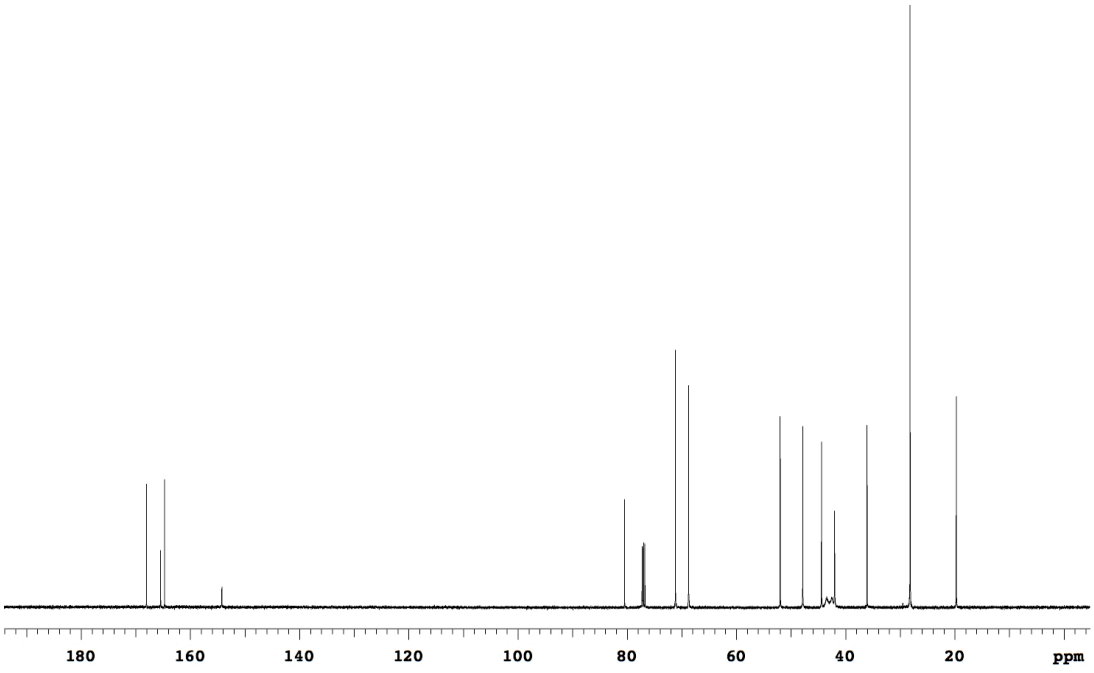
Compound **6b** was prepared from (0.15 g, 0.3 mmol) of cyclo(Thr-Gly) derivative **4b**. Flash chromatography (2 % to 5 % MeOH/CH_2Cl_2) afforded 0.11 g (94 %) **6b** as a white solid.

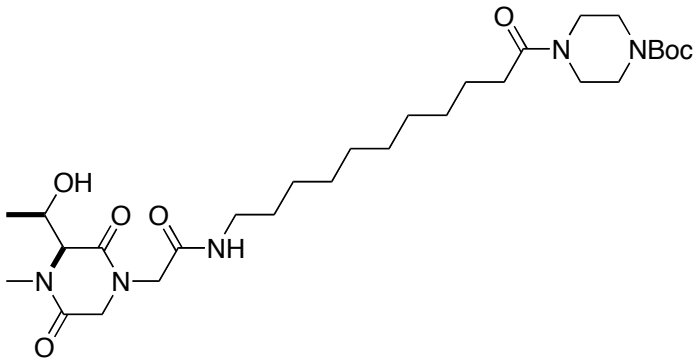
^{1}H NMR (500 MHz, $CDCl_3$) δ 4.93 (br, 1H), 4.85 (d, 1H, J = 16.1 Hz), 4.42-4.32 (m, 2H), 3.77 (d, 1H, J = 1.7 Hz), 3.66 (d, 1H, J = 16.1 Hz), 3.62-3.32 (m, 8H), 3.06 (s, 3H), 1.43 (s, 9H), 1.27 (d, 3H, J = 6.5 Hz)

^{13}C NMR (125 MHz, $CDCl_3$) δ 168.0, 165.4, 164.7, 154.2, 80.5, 71.2, 68.7, 52.0, 47.8, 44.4, 42.0, 36.1, 28.2, 19.7

MS (ESI, m/z) 399.2 $(M+H)^+$


8
7
6
5
4
3
2
1
ppm


180
160
140
120
100
80
60
40
20
ppm



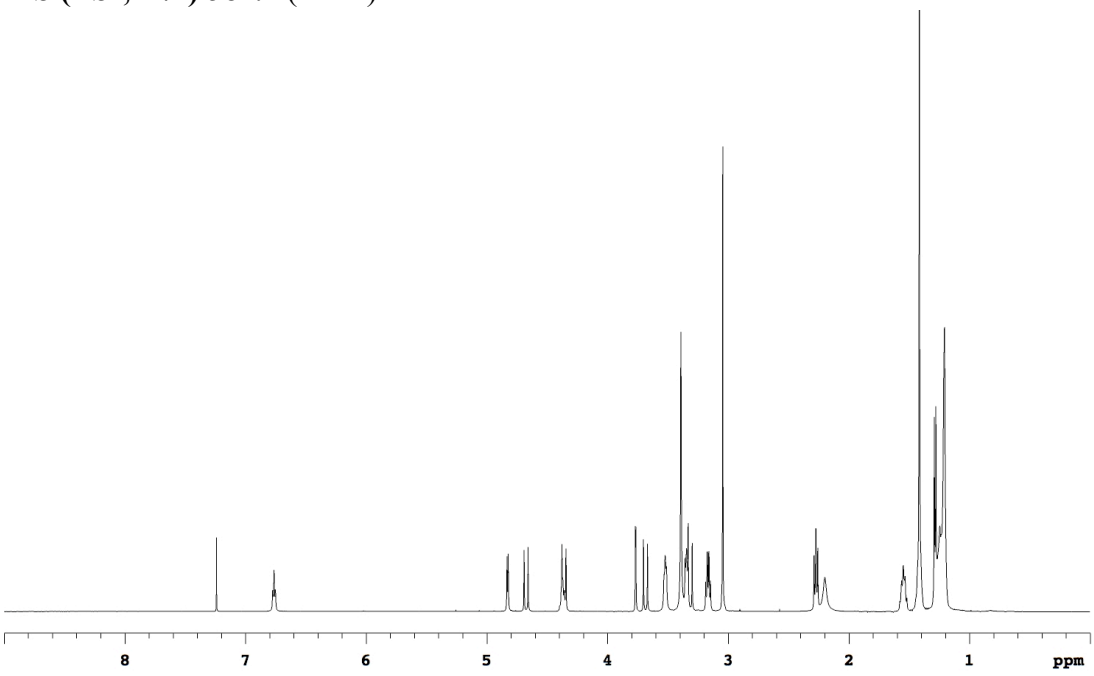
7b

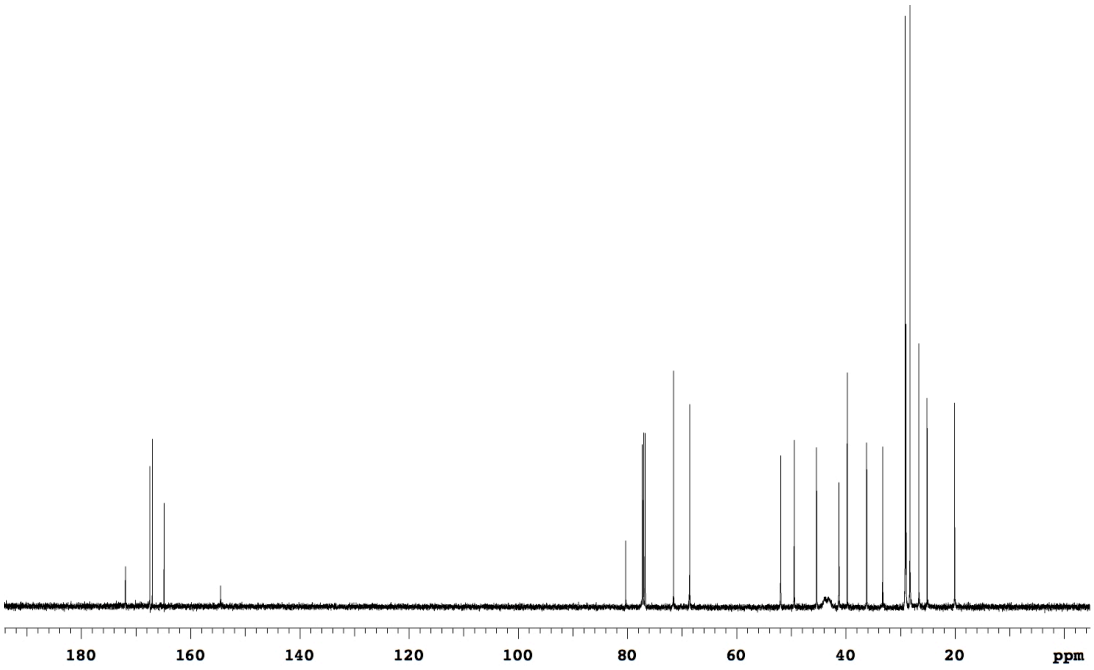
Compound **7b** was prepared from (0.18 g, 0.3 mmol) of Thr-Gly derivative **5b**. Flash chromatography (2 % to 10 % MeOH/CH_2Cl_2) afforded 0.13 g (83 %) **7b** as a white solid.

^{1}H NMR (500 MHz, $CDCl_3$) δ 6.76 (t, 1H, J = 5.6 Hz), 4.83 (d, 1H, J = 4.8 Hz), 4.67 (d, 1H, J = 16.3 Hz), 4.40-4.33 (m, 2H), 3.77 (d, 1H, J = 1.9 Hz), 3.68 (d, 1H, J = 16.3 Hz), 3.55-3.49 (m, 2H), 3.42-3.32 (m, 6H), 3.17 (q, 2H, J = 6.7 Hz), 3.04 (s, 3H), 2.31-2.15 (m, 4H), 1.59-1.51 (m, 2H), 1.42 (s, 9H), 1.32-1.17 (m, 15H)

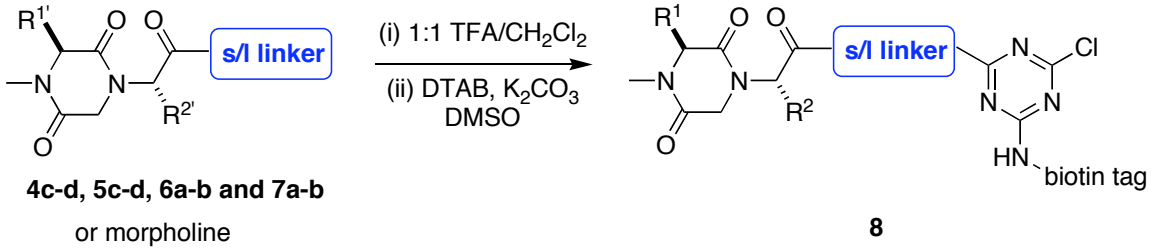
^{13}C NMR (125 MHz, $CDCl_3$) δ 171.9, 167.4, 167.0, 164.9, 154.5, 80.3, 71.6, 68.6, 51.9, 49.4, 45.3, 41.2, 39.7, 36.2, 33.2, 29.1, 29.1, 29.0, 29.0, 28.3, 26.6, 25.1, 20.1

MS (ESI, m/z) 582.4 $(M+H)^+$





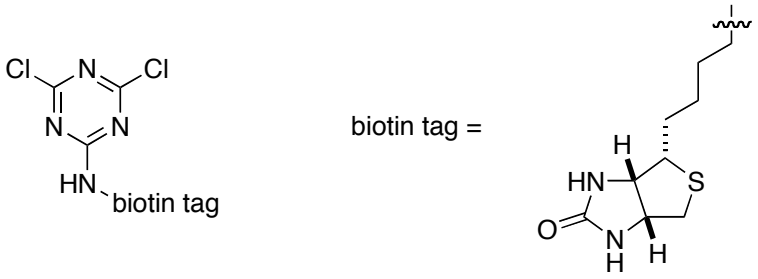
General Procedure for Preparation of Biotin Labeled Monomers 9. Boc-protected monomeric compounds **4c-d**, **5c-d**, **6a-b** and **7a-b** (1.0 equiv) were treated with 50% TFA in CH_2Cl_2, stirred in sealed glass vials for 10 h and solvent was removed. The resulting residue was dissolved in DMSO (0.033 M) and biotin tag derivative DTAB (1.0 equiv) and K_2CO_3 (4.0 equiv) were added. The suspension was stirred for 2 h and the crude product **8** was used for the next step without further purification. Analytical HPLC and MALDI MS were used to check the purities of these crude samples.



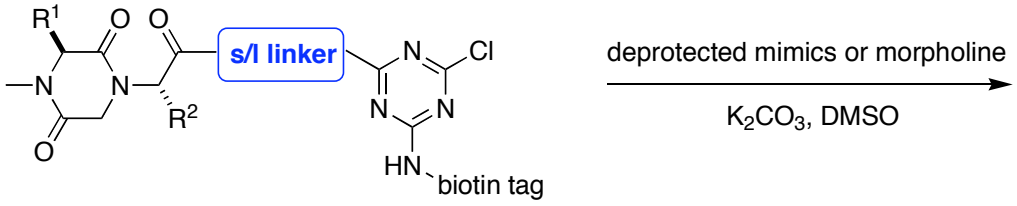
4c-d, 5c-d, 6a-b and 7a-b

or morpholine

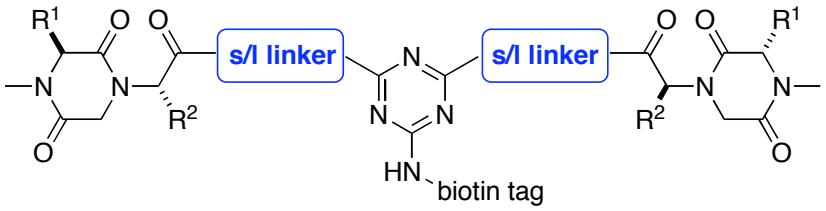
8



Preparation of Biotin Labeled Dimers 9 and Monomers 9ai-ii. Dimerization was carried out in 2 ml glass vials. The monomeric compounds **4c-d**, **5c-d**, **6a-b** and **7a-b** were treated with 50% TFA/CH_2Cl_2 for 10 h. Stock solutions of deprotected mimics (0.03 M) or morpholine (0.33 M) in DMSO were prepared. One stock solution of deprotected **74** (0.5 ml) or morpholine (0.5 ml) were added to a glass vial containing pre-formed **76** (0.03 M, 0.5 ml) and followed with solid K_2CO_3 (4 equiv). The reaction vial was sealed and the reaction mixture was stirred at 25 °C for 96 h then the DMSO was lyophilized. Aqueous HCl solution (5%, *ca* 1-2 ml) was added to the above solid residue and sonicated for 2 min. Most of the compounds were dissolved in the acidic aqueous solution and all K_2CO_3 was converted to KCl. The aqueous solution was dried by carefully blowing N_2. The solid residue was redissolved in dry methanol and stirred for 2 min. Most KCl was precipitated from the solution and removed by filtration. The MeOH solution was then concentrated to give crude bivalent compounds **9**. All crude products were purified by preprative HPLC and analyzed by analytical HPLC (10-90% B in 30 min), and also by ^{1}H NMR, gCOSY (21 out of 45 compounds) and MALDI-MS.

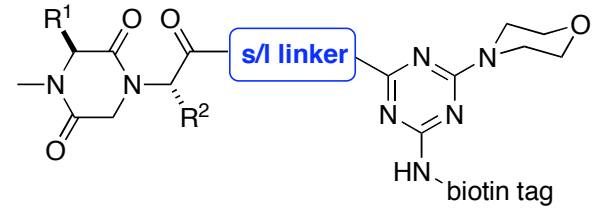


8

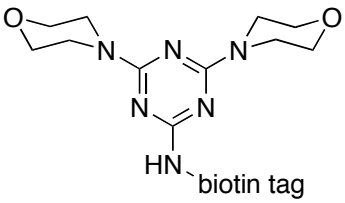


9

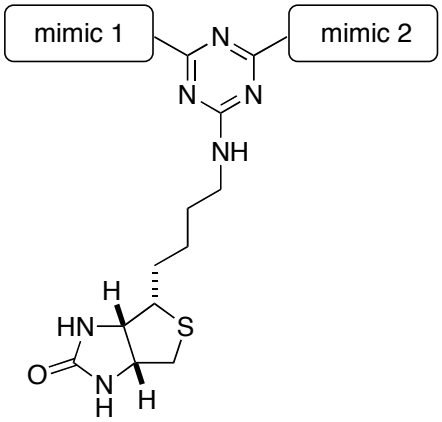
or



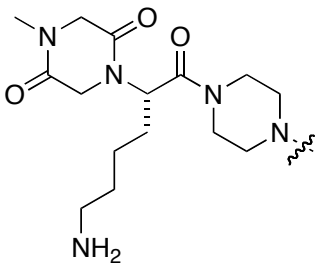
or



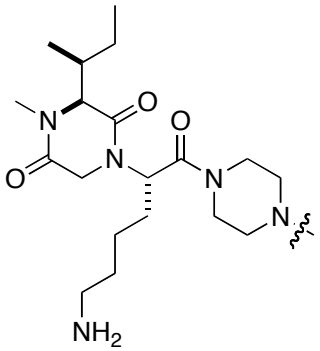
9ai-hi **9ii**



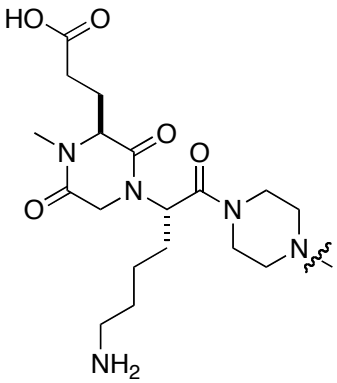
General structure of 9



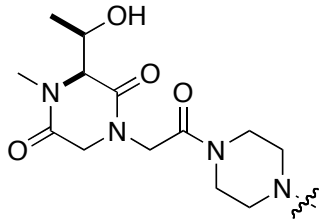
a



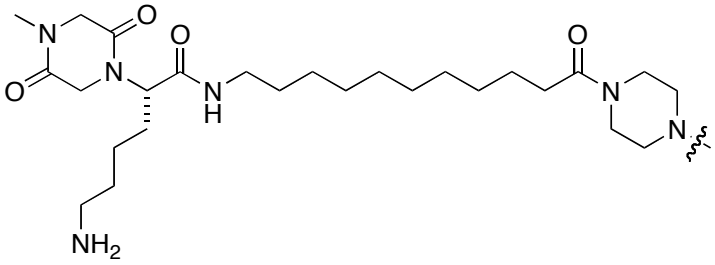
b



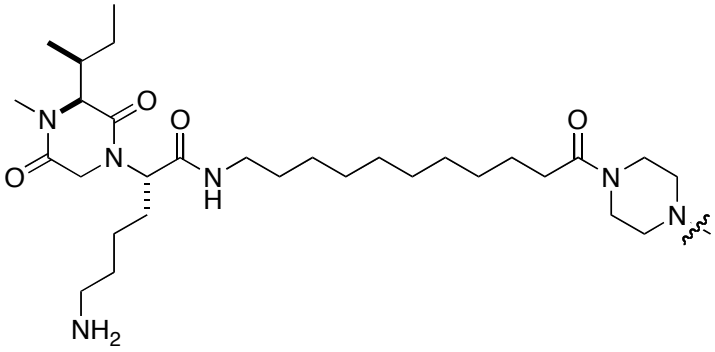
c



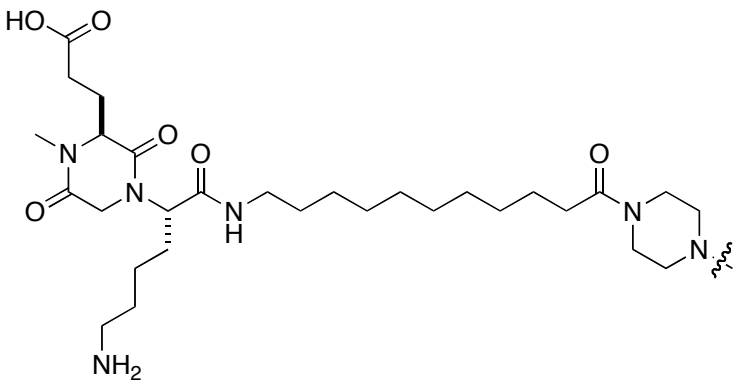
d



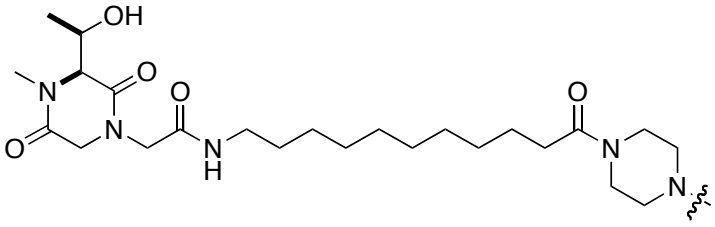
e



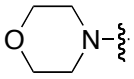
f



g



h



i (cap)

UV purity (%) of crude product 9

	i	a	b	c	d	e	f	g	h
i(cap)	100								
a	100	75							
b	98	71	89						
c	100	72	68	80					
d	100	47	80	58	89				
e	95	45	59	43	71	64			
f	100	88	87	70	79	56	100		
g	98	93	92	70	58	41	72	86	
h	100	51	57	58	81	65	76	36	85

UV purities (%)
below 70
70 – 85
85 and up

UV purity (%) of compounds 9 after purification

	i	a	b	c	d	e	f	g	h
i(cap)	100								
a	100	100							
b	100	98	100						
c	100	100	100	100					
d	100	100	100	100	100				
e	100	100	100	100	100	100			
f	100	100	100	100	100	100	100		
g	100	100	100	100	100	100	100	100	
h	100	100	100	100	97	100	100	100	100

UV purities (%)
below 85
85 – 90
90 and up

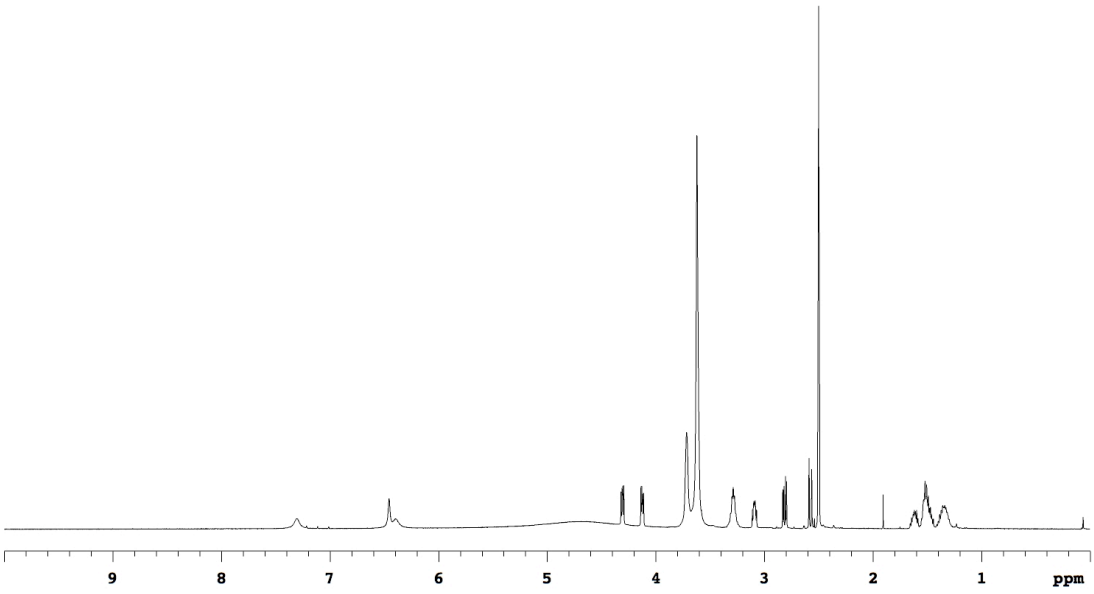
Mass of bivalent compounds 9

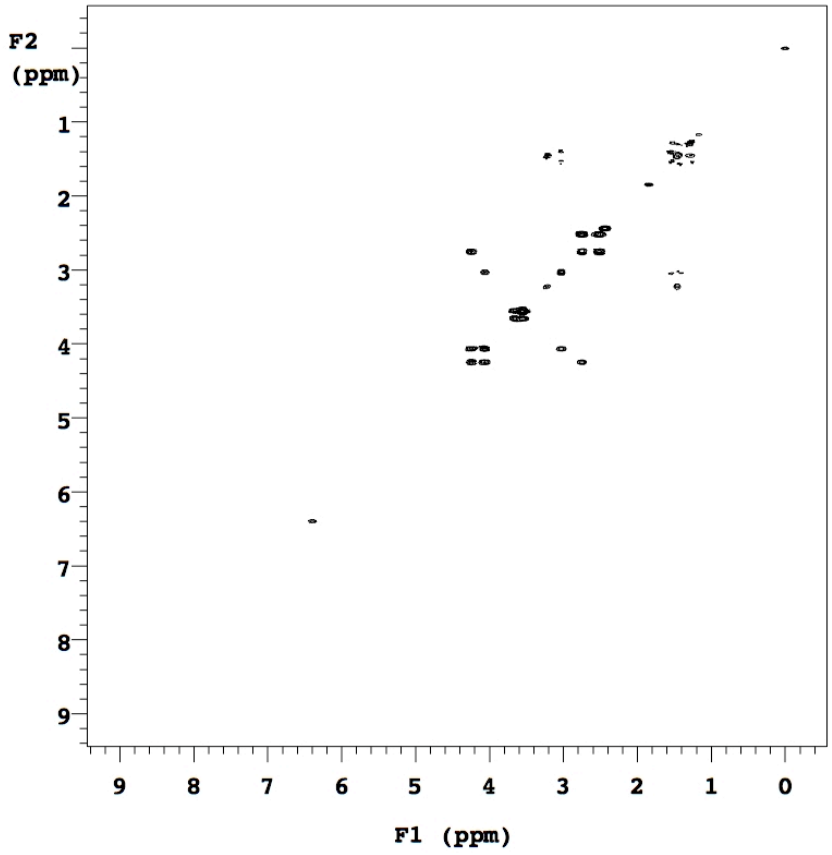
	MALDI-MS			MALDI-MS	
Comp'd **9**	Calc'd $(M+H)^+$	Found $(M+H)^+$	Comp'd **9**	Calc'd $(M+H)^+$	Found $(M+H)^+$
ii	465.2	465.0	**fb**	1236.8	1236.9
ai	703.4	703.3	**fc**	1252.8	1252.9
aa	941.5	941.5	**fd**	1153.7	1153.7
bi	759.4	759.4	**fe**	1363.9	1364.1
ba	997.6	997.6	**ff**	1420.0	1420.0
bb	1053.6	1053.7	**gi**	958.6	958.7
ci	775.4	775.3	**ga**	1196.7	1196.8
ca	1013.5	1013.5	**gb**	1252.8	1252.8
cb	1069.6	1069.5	**gc**	1268.7	1268.8
cc	1085.6	1085.5	**gd**	1169.7	1169.8
di	676.3	676.3	**ge**	1379.9	1380.0
da	914.5	914.3	**gf**	1435.9	1436.0
db	970.5	970.5	**gg**	1451.9	1452.0
dc	986.5	986.4	**hi**	859.5	859.4
dd	887.4	887.4	**ha**	1097.6	1097.7

ei	886.5	886.6	**hb**	1153.7	1153.8
ea	1124.7	1124.7	**hc**	1169.7	1169.7
eb	1180.7	1181.0	**hd**	1070.6	1070.7
ec	1196.7	1196.8	**he**	1280.8	1280.8
ed	1097.6	1097.7	**hf**	1336.9	1336.9
ee	1307.8	1308.1	**hg**	1352.8	1352.9
fi	942.6	942.7	**hh**	1253.7	1253.7
fa	1180.7	1180.8			

Data for 9ii

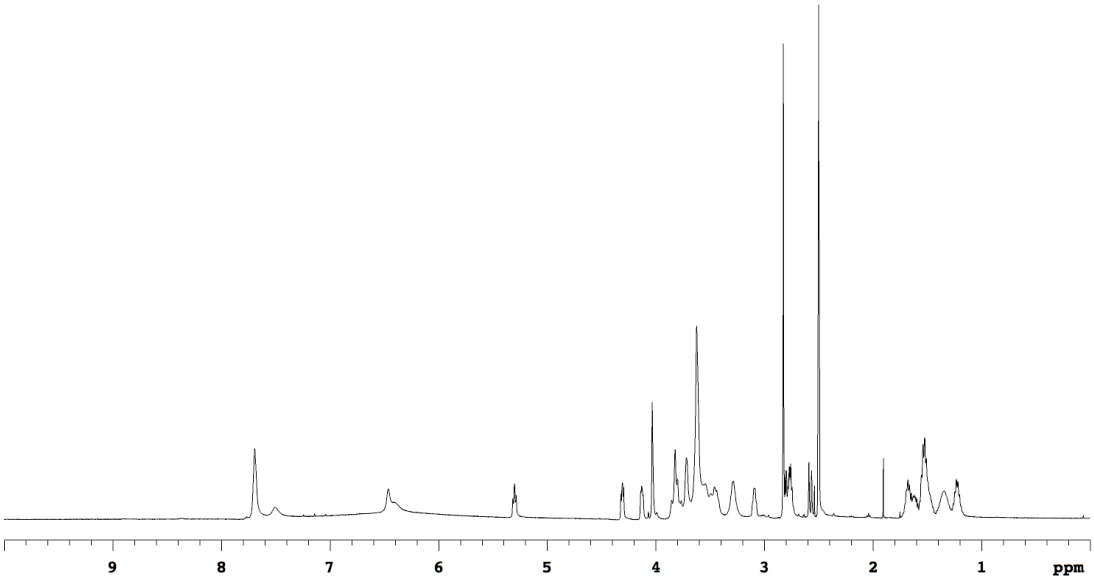
^{1}H NMR (500 MHz, DMSO-d_6) δ 7.30 (br, 1H), 6.46 (br, 1H), 6.40 (br, 1H), 4.31 (dd, 1H, J = 5.2, 7.8 Hz), 4.12 (dd, 1H, J = 4.5, 7.8 Hz), 3.72 (br, 4H), 3.62 (br, 12H), 3.35-3.22 (m, 2H), 3.12-3.06 (m, 1H), 2.81 (dd, 1H, J = 5.2, 12.6 Hz), 2.58 (d, 1H, J = 12.6 Hz), 1.67-1.58 (m, 1H), 1.57-1.43 (m, 3H), 1.42-1.27 (m, 2H)

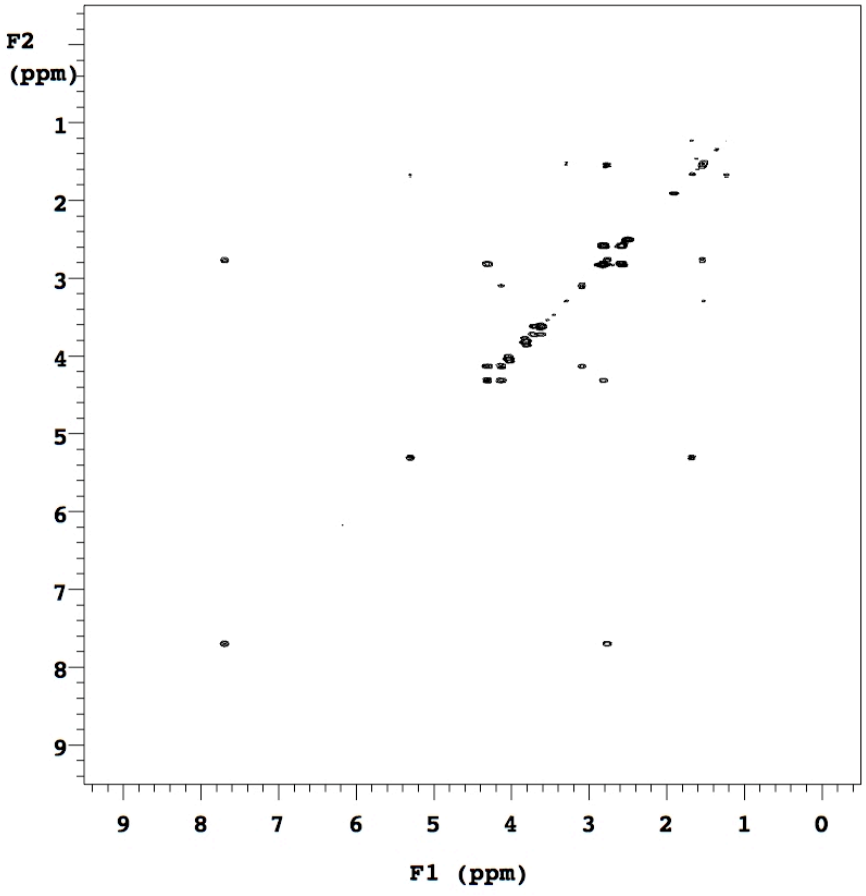




Data for 9ai

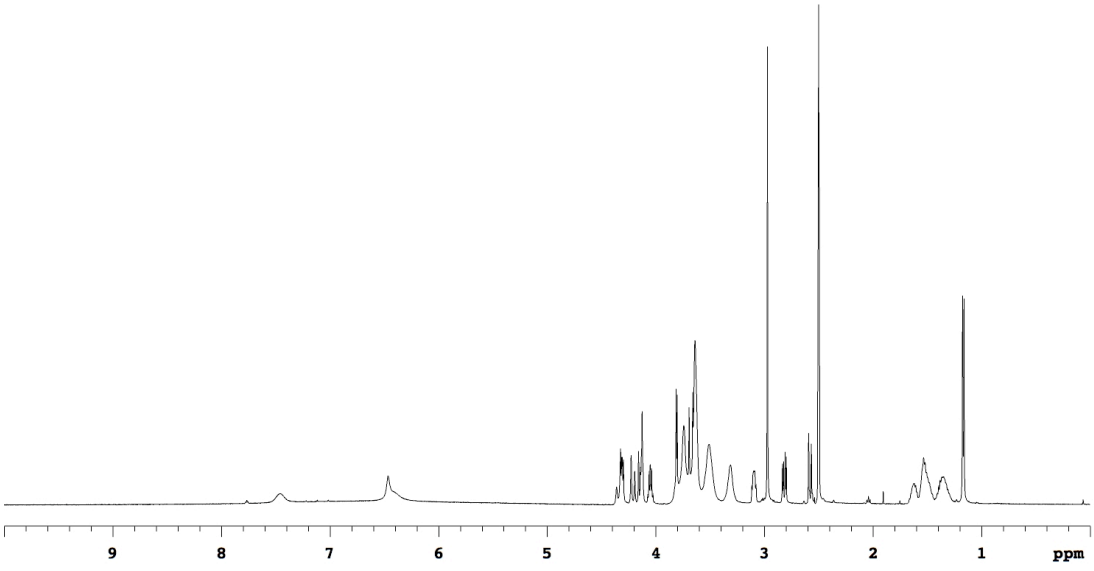
^{1}H NMR (500 MHz, DMSO-d_6) δ 7.69 (br, 2H), 7.50 (br, 1H), 6.46 (br, 1H), 6.39 (br, 1H), 5.30 (t, 1H, J = 7.4 Hz), 4.34-4.27 (m, 1H), 4.16-4.09 (m, 1H), 4.03 (br, 2H), 3.89-3.75 (m, 3H), 3.72 (br, 2H), 3.68-3.38 (m, 13H), 3.29 (br, 2h), 3.13-3.05 (m, 1H), 2.85-2.71 (m, 6H), 2.58 (d, 1H, J = 12.5 Hz), 1.75-1.43 (m, 8H), 1.43-1.28 (m, 2H), 1.28-1.15 (m, 2H)

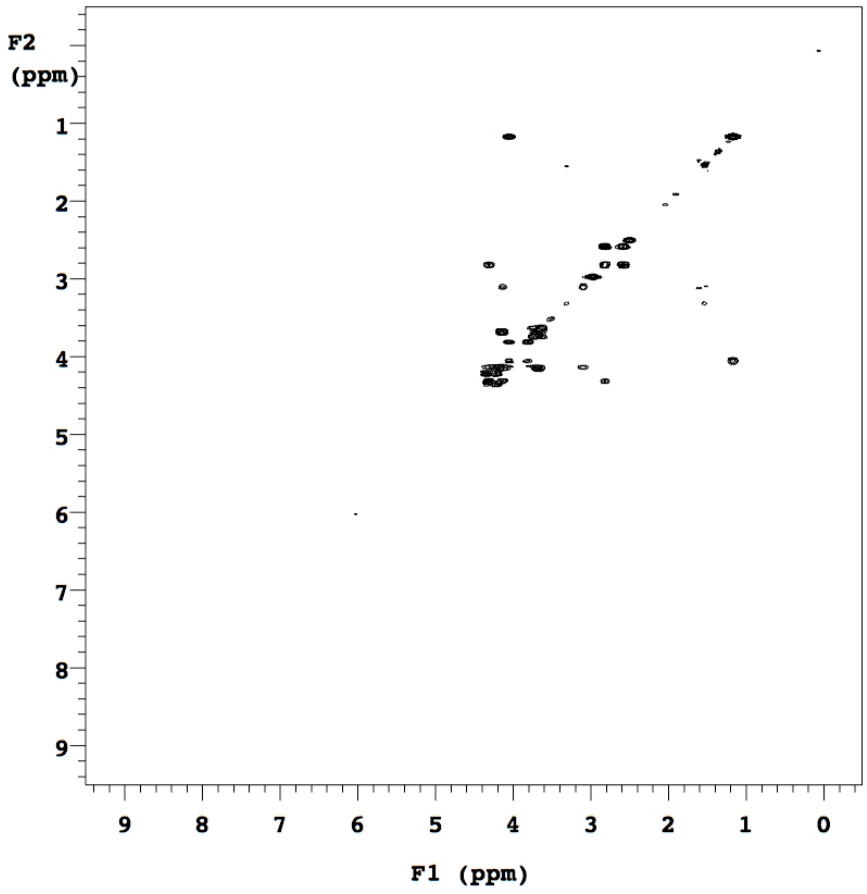

9
8
7
6
5
4
3
2
1
ppm


F2
(ppm)
1
2
3
4
5
6
7
8
9
9
8
7
6
5
4
3
2
1
0
F1 (ppm)

Data for 9di

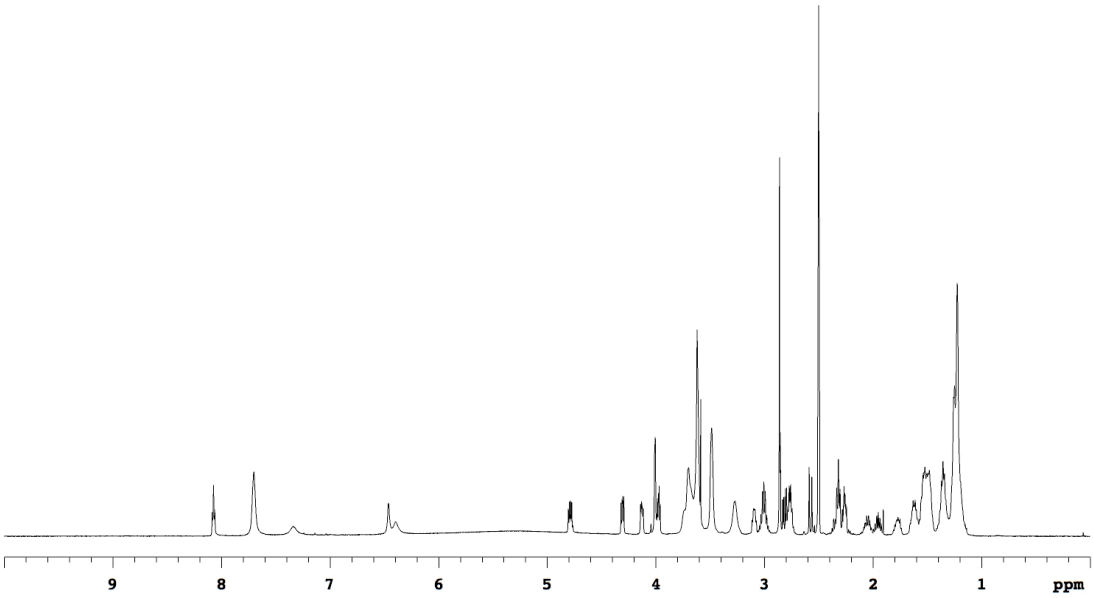
^{1}H NMR (500 MHz, DMSO-d_6) δ 7.46 (br, 1H), 6.47 (br, 2H), 4.38-4.28 (m, 2H), 4.21 (d, 1H, J = 16.5 Hz), 4.18-4.10 (m, 2H), 4.08-4.02 (m, 1H), 3.87-3.59 (m, 12H), 3.51 (br, 4H), 3.31 (br, 2h), 3.13-3.06 (m, 2H), 2.97 (s, 3H), 2.82 (dd, 1H, J = 5.0, 12.5 Hz), 2.58 (d, 1H, J = 12.5 Hz), 1.68-1.58 (m, 1H), 1.58-1.43 (m, 3H), 1.43-1.26 (m, 2H), 1.17 (d, 3H, J = 6.5 Hz)

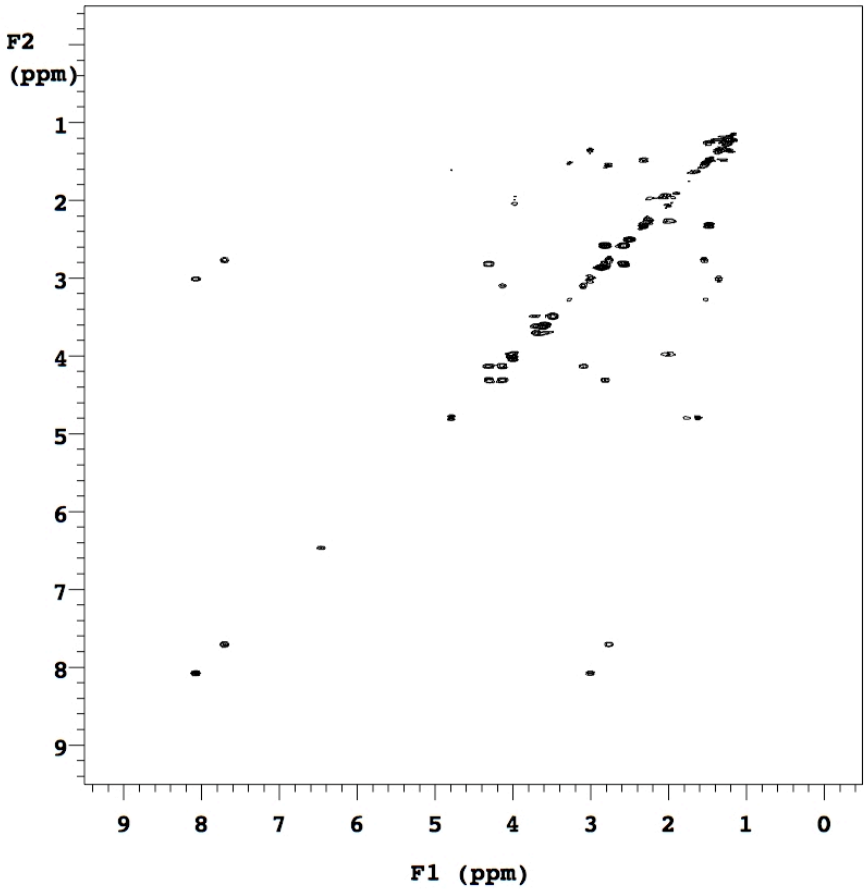




Data for 9gi

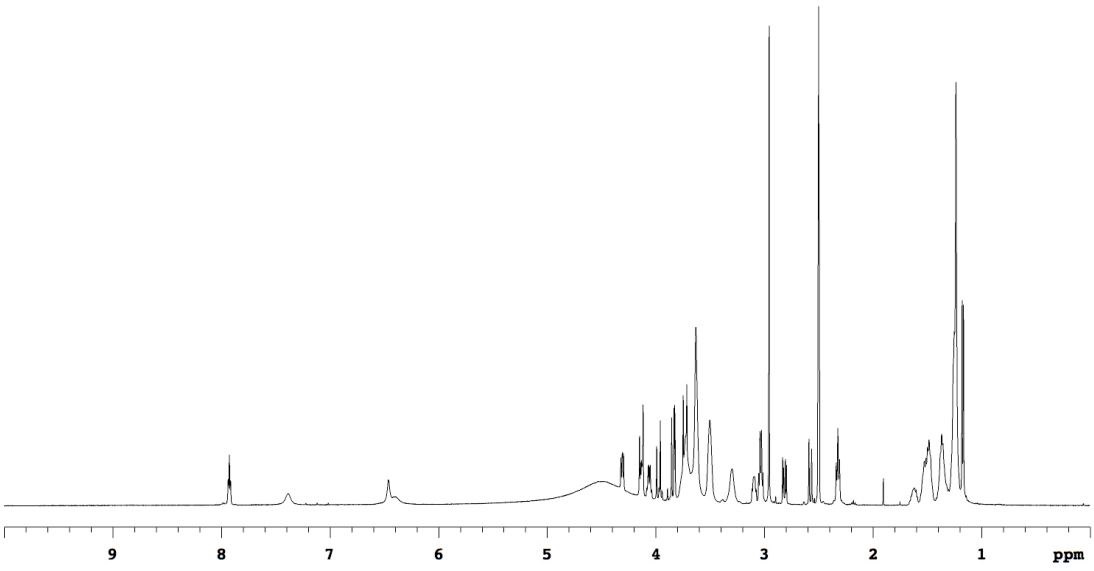
^{1}H NMR (500 MHz, DMSO-d_6) δ 8.07 (t, 1H, J = 5.7 Hz), 7.70 (br, 2H), 7.33 (br, 1H), 6.46 (br, 1H), 6.39 (br, 1H), 4.84-4.74 (m, 1H), 4.34-4.28 (m, 1H), 4.16-4.10 (m, 1H), 4.04-3.94 (m, 3H), 3.79-3.56 (m, 12H), 3.49 (br, 4H), 3.27 (br, 2h), 3.13-3.06 (m, 1H), 3.06-2.95 (m, 2H), 2.86 (s, 3H), 2.84-2.71 (m, 2H), 2.57 (d, 1H, J = 12.4 Hz), 2.39-2.23 (m, 3H), 2.10-1.90 (m, 2H), 1.83-1.73 (m, 1H), 1.68-1.58 (m, 2H), 1.58-1.43 (m, 8H), 1.43-1.30 (m, 4H), 1.30-1.12 (m, 16H)

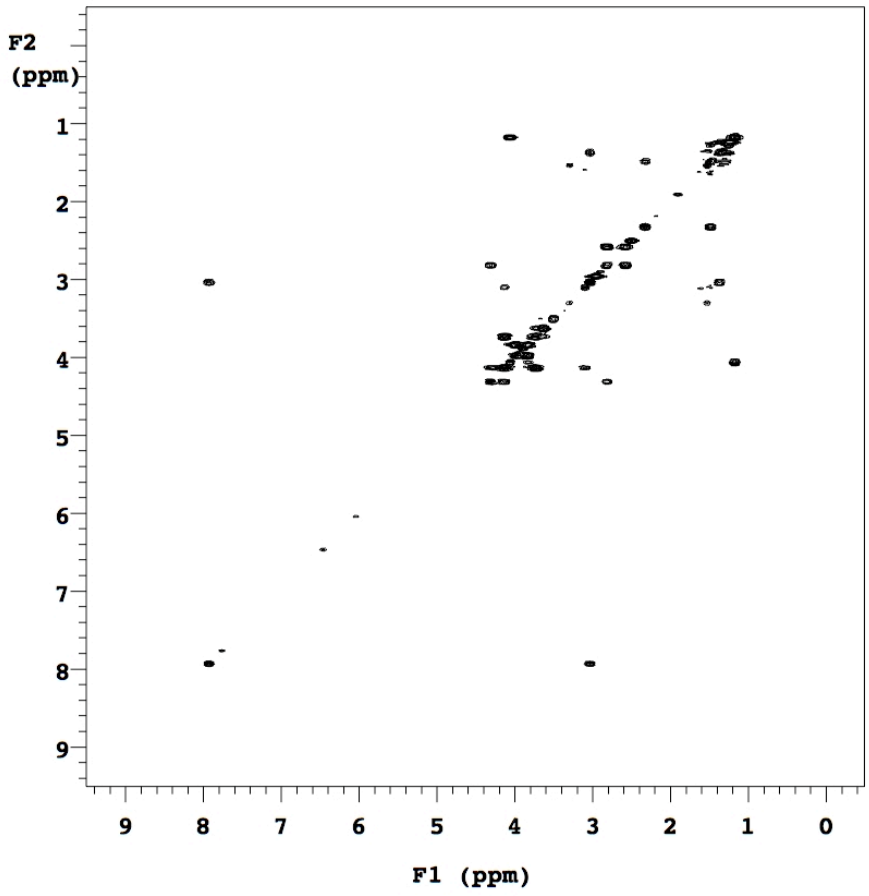

9
8
7
6
5
4
3
2
1
ppm


F2
(ppm)
1
2
3
4
5
6
7
8
9
9
8
7
6
5
4
3
2
1
0
F1 (ppm)

Data for 9hi

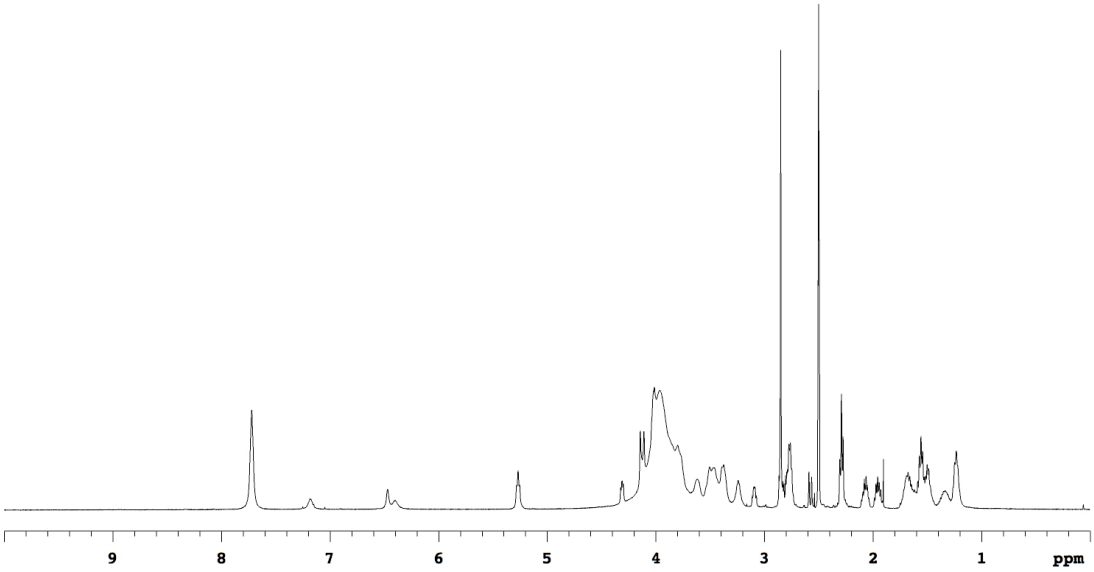
^{1}H NMR (500 MHz, DMSO-d_6) δ 7.93 (t, 1H, J = 5.5 Hz), 7.39 (br, 1H), 6.46 (br, 1H), 6.39 (br, 1H), 4.34-4.28 (m, 1H), 4.17-4.10 (m, 2H), 4.09-4.03 (m, 1H), 3.97 (d, 1H, J = 16.4 Hz), 3.88-3.81 (m, 2H), 3.79-3.56 (m, 13H), 3.50 (br, 4H), 3.30 (br, 2h), 3.12-3.07 (m, 2H), 3.06-3.00 (m, 2H), 2.96 (s, 3H), 2.81 (dd, 1H, J = 5.2, 12.6 Hz), 2.58 (d, 1H, J = 12.6 Hz), 2.32 (t, 2H, J = 7.4 Hz), 1.68-1.58 (m, 1H), 1.57-1.44 (m, 5H), 1.42-1.31 (m, 4H), 1.30-1.20 (m, 12H), 1.17 (d, 3H, J = 6.6 Hz)

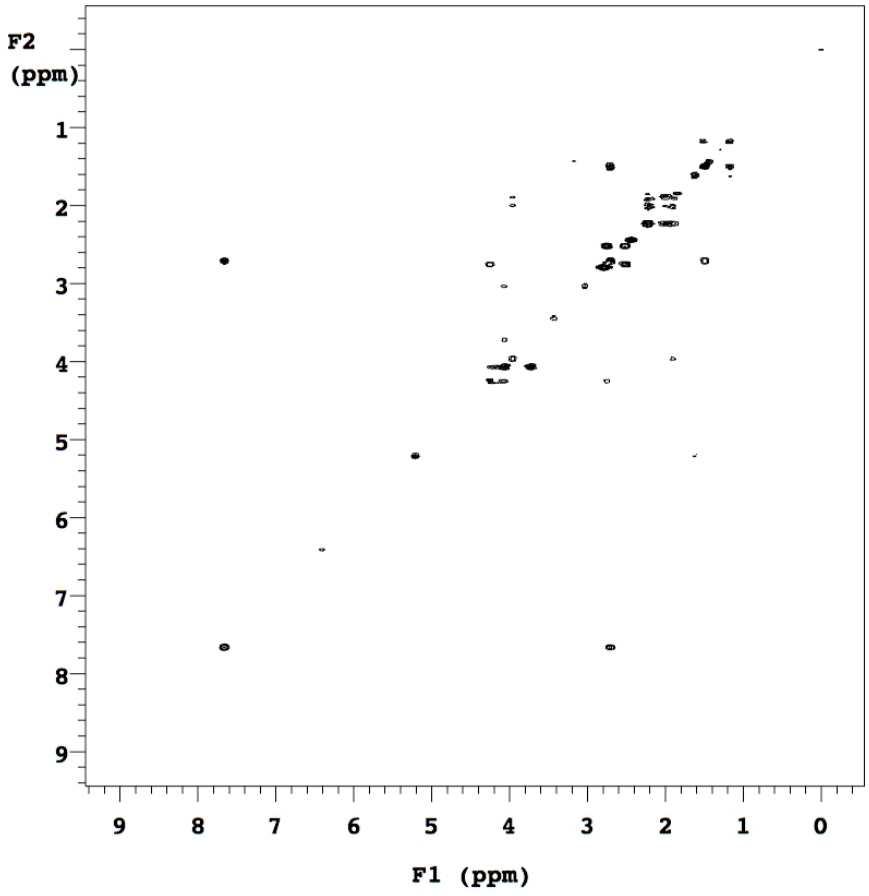




Data for 9cc

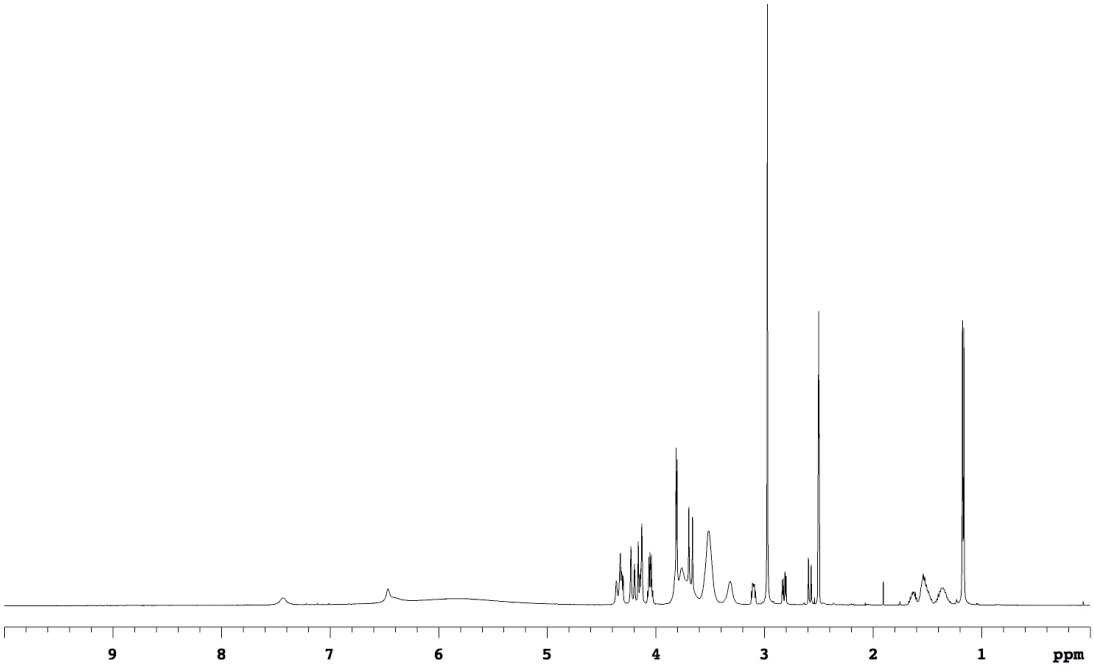
^{1}H NMR (500 MHz, DMSO-d_6) δ 7.72 (br, 4H), 7.18 (br, 1H), 6.47 (br, 1H), 6.40 (br, 1H), 5.27 (t, 2H, J = 7.2 Hz), 4.40-3.69 (m, 15H), 3.69-3.56 (m, 2H), 3.56-3.30 (m, 8H), 3.24 (br, 2H), 3.13-3.06 (m, 1H), 2.90-2.70 (m, 10H), 2.58 (d, 1H, J = 12.5 Hz), 2.29 (t, 4H, J = 7.7 Hz), 2.14-2.01 (m, 2H), 2.01-1.89 (m, 2H), 1.78-1.43 (m, 12H), 1.42-1.29 (m, 2H), 1.29-1.15 (m, 4H)

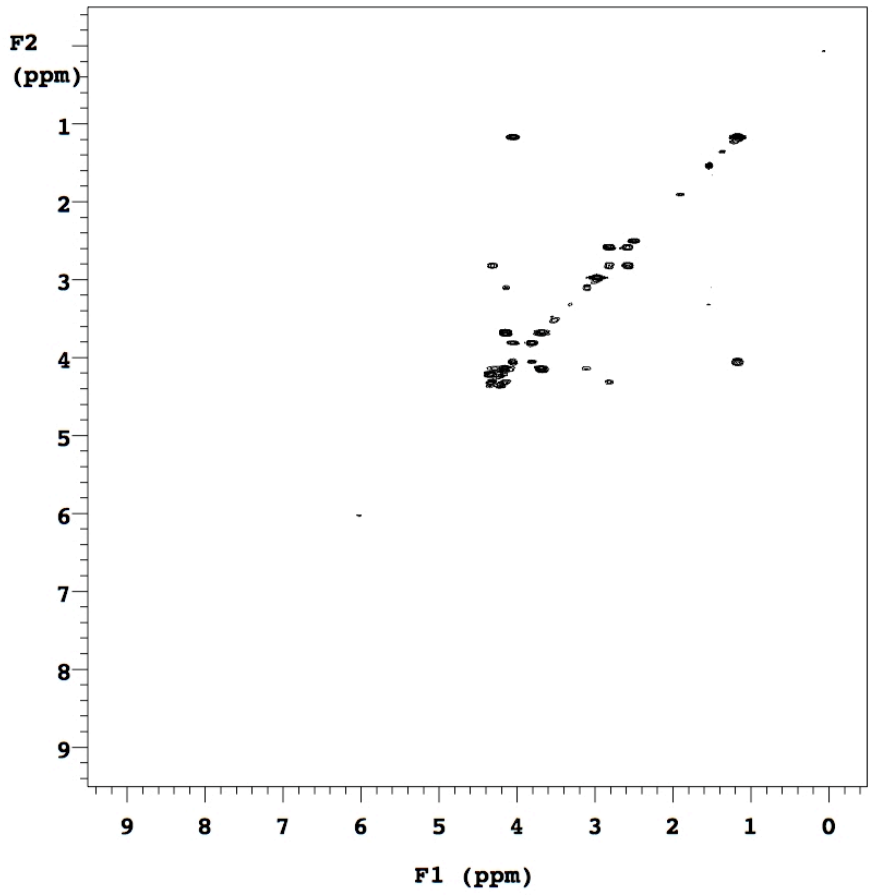

9
8
7
6
5
4
3
2
1
ppm


F2
(ppm)
1
2
3
4
5
6
7
8
9
9
8
7
6
5
4
3
2
1
0
F1 (ppm)

Data for 9dd

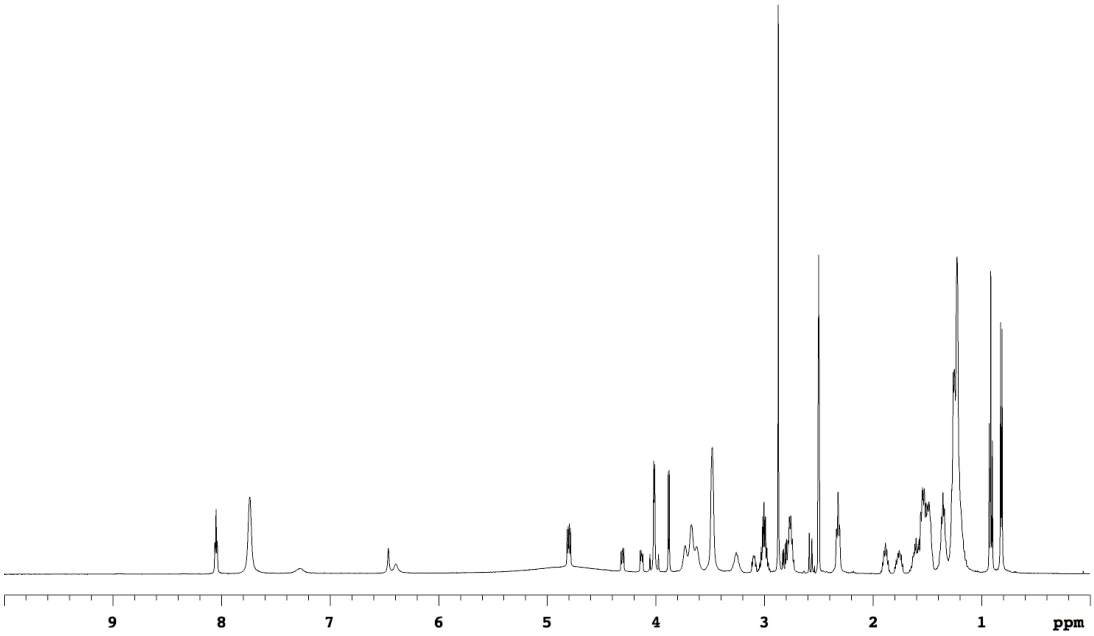
^{1}H NMR (500 MHz, DMSO-d_6) δ 7.43 (br, 1H), 6.47 (br, 2H), 4.40-4.28 (m, 3H), 4.21 (d, 2H, J = 16.5 Hz), 4.18-4.10 (m, 3H), 4.09-4.01 (m, 2H), 3.88-3.60 (m, 14H), 3.51 (br, 8H), 3.32 (br, 2H), 3.13-3.07 (m, 1H), 2.97 (s, 6H), 2.82 (dd, 1H, J = 5.0, 12.5 Hz), 2.58 (d, 1H, J = 12.5 Hz), 1.69-1.59 (m, 1H), 1.59-1.44 (m, 3H), 1.44-1.28 (m, 2H), 1.17 (d, 6H, J = 6.5 Hz)

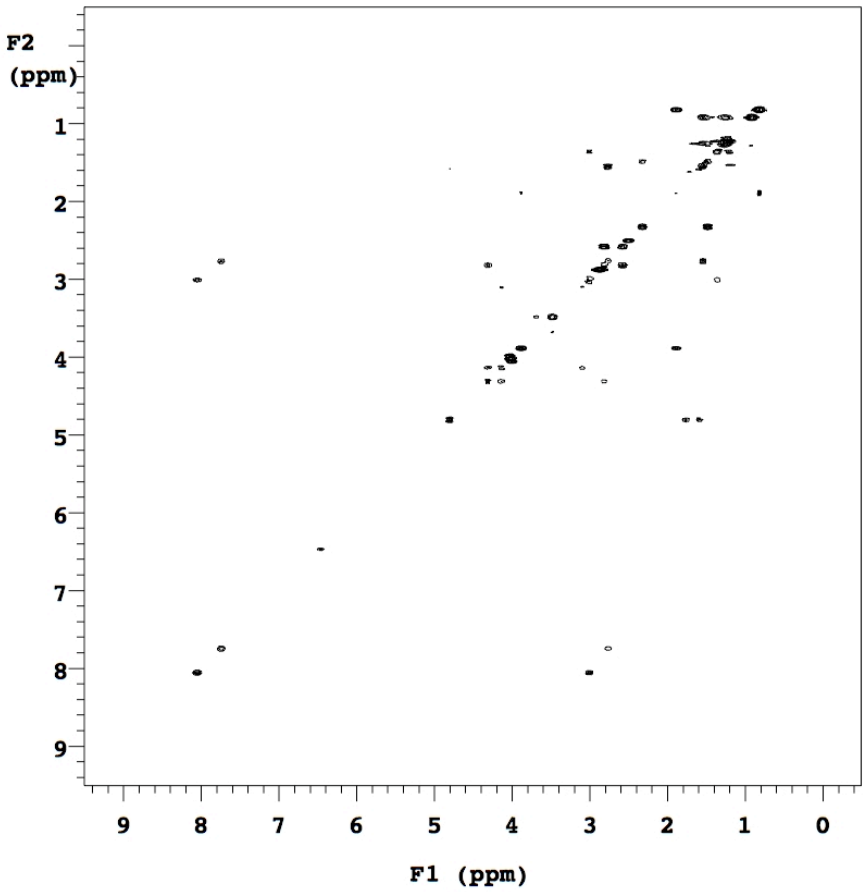




Data for 9ff

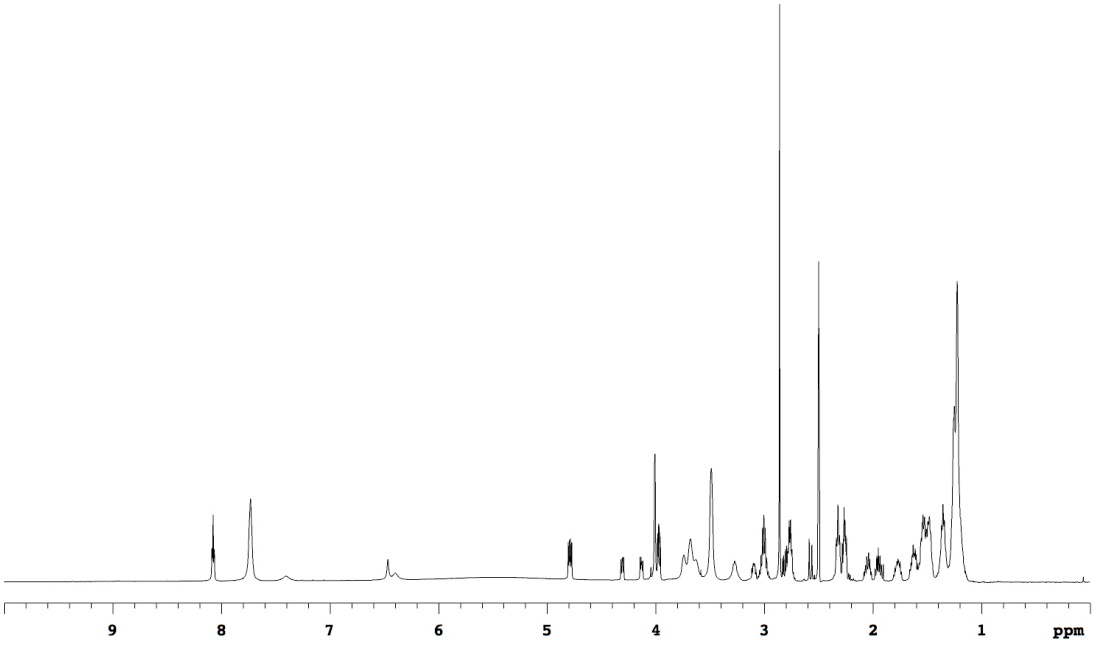
^{1}H NMR (500 MHz, DMSO-d_6) δ 8.05 (t, 2H, J = 5.6 Hz), 7.74 (br, 4H), 7.28 (br, 1H), 6.46 (br, 1H), 6.39 (br, 1H), 4.84-4.77 (m, 2H), 4.31 (dd, 1H, J = 5.3, 7.8 Hz), 4.13 (dd, 1H, J = 4.5, 7.8 Hz), 4.07-3.95 (m, 4H), 3.88 (d, 2H, J = 4.9 Hz), 3.78-3.57 (m, 8H), 3.48 (br, 8H), 3.32-3.19 (m, 2H), 3.13-3.07 (m, 1H), 3.06-2.94 (m, 4H), 2.87 (s, 6H), 2.85-2.70 (m, 5H), 2.57 (d, 1H, J = 12.5 Hz), 2.32 (t, 4H, J = 7.4 Hz), 1.93-1.84 (m, 2H), 1.81-1.71 (m, 2H), 1.67-1.43 (m, 16H), 1.41-1.31 (m, 6H), 1.31-1.12 (m, 30H), 0.92 (t, 6H, J = 7.4 Hz), 0.82 (d, 6H, J = 6.8 Hz)

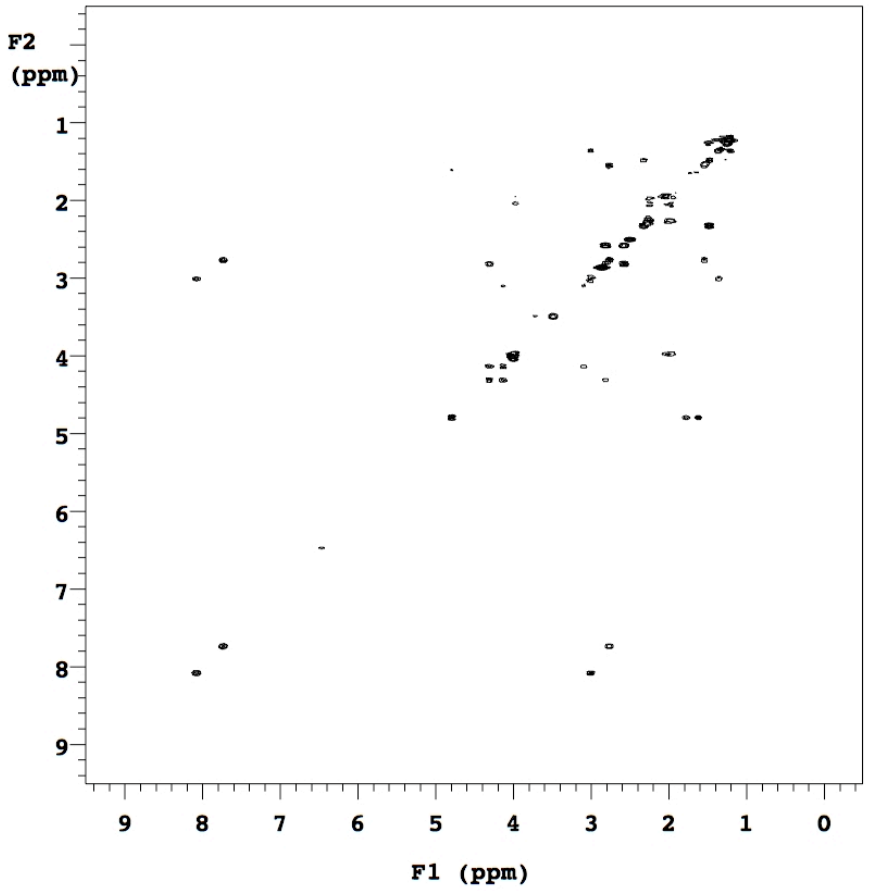

9
8
7
6
5
4
3
2
1
ppm


F2
(ppm)
1
2
3
4
5
6
7
8
9
9
8
7
6
5
4
3
2
1
0
F1 (ppm)

Data for 9gg

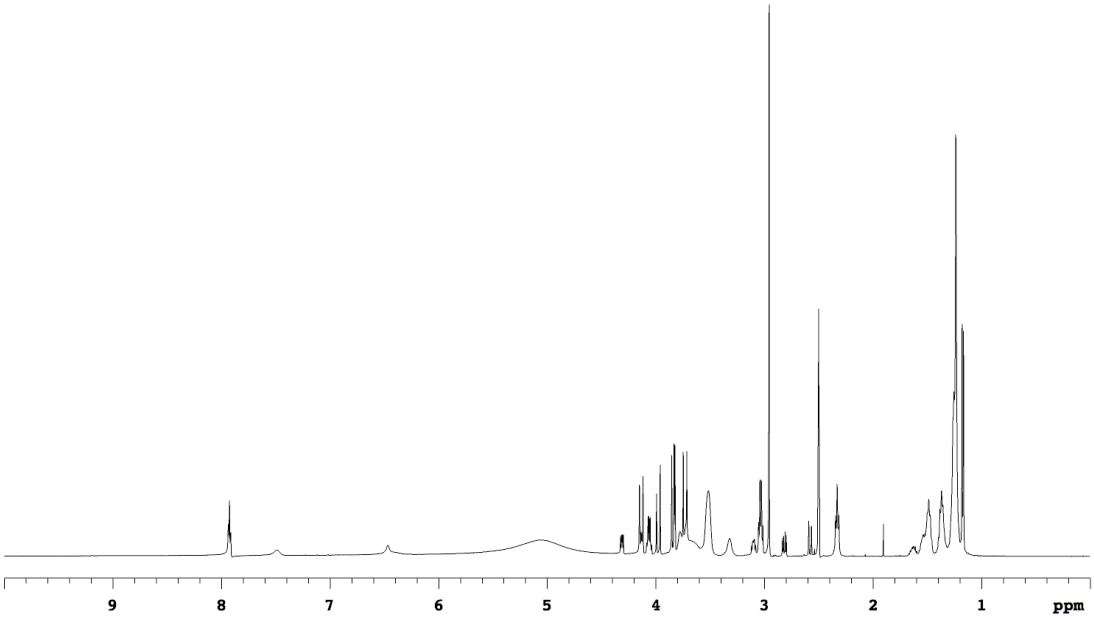
^{1}H NMR (500 MHz, DMSO-d_6) δ 8.08 (t, 2H, J = 5.6 Hz), 7.73 (br, 4H), 7.40 (br, 1H), 6.47 (br, 1H), 6.40 (br, 1H), 4.83-4.76 (m, 2H), 4.31 (dd, 1H, J = 5.0, 7.7 Hz), 4.13 (dd, 1H, J = 4.5, 7.7 Hz), 4.06-3.94 (m, 6H), 3.80-3.57 (m, 8H), 3.49 (br, 8H), 3.34-3.21 (m, 2H), 3.14-3.07 (m, 1H), 3.07-2.94 (m, 4H), 2.86 (s, 6H), 2.84-2.71 (m, 5H), 2.58 (d, 1H, J = 12.5 Hz), 2.39-2.20 (m, 8H), 2.10-2.00 (m, 2H), 2.00-1.89 (m, 2H), 1.83-1.72 (m, 2H), 1.68-1.58 (m, 3H), 1.58-1.43 (m, 11H), 1.42-1.31 (m, 8H), 1.31-1.11 (m, 26H)

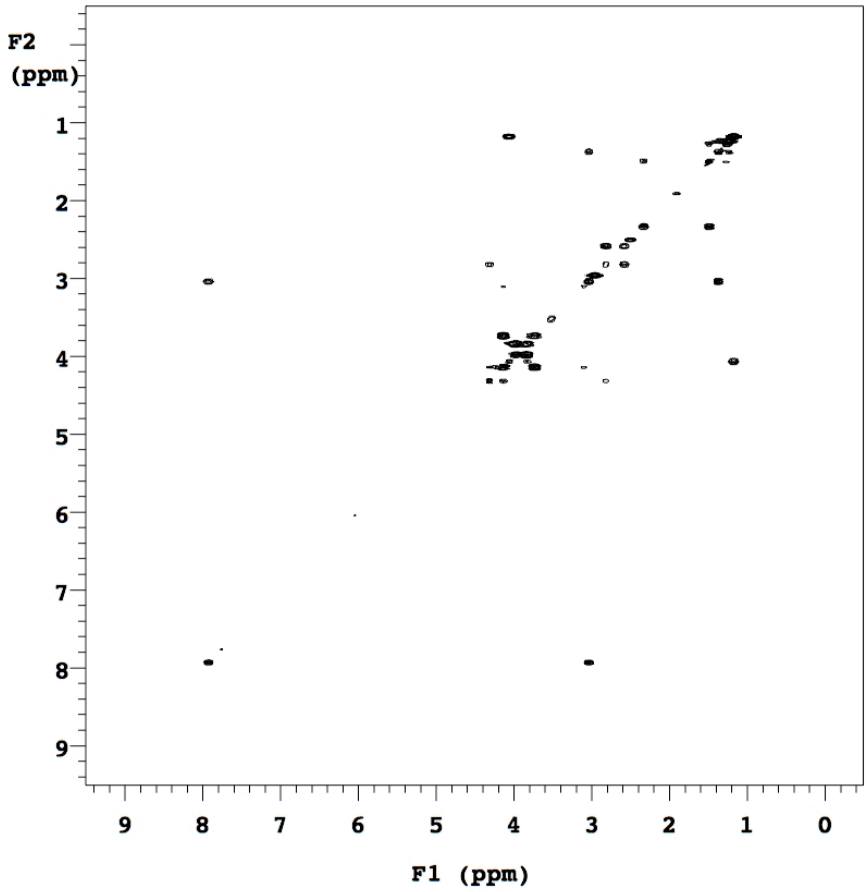




Data for 9hh

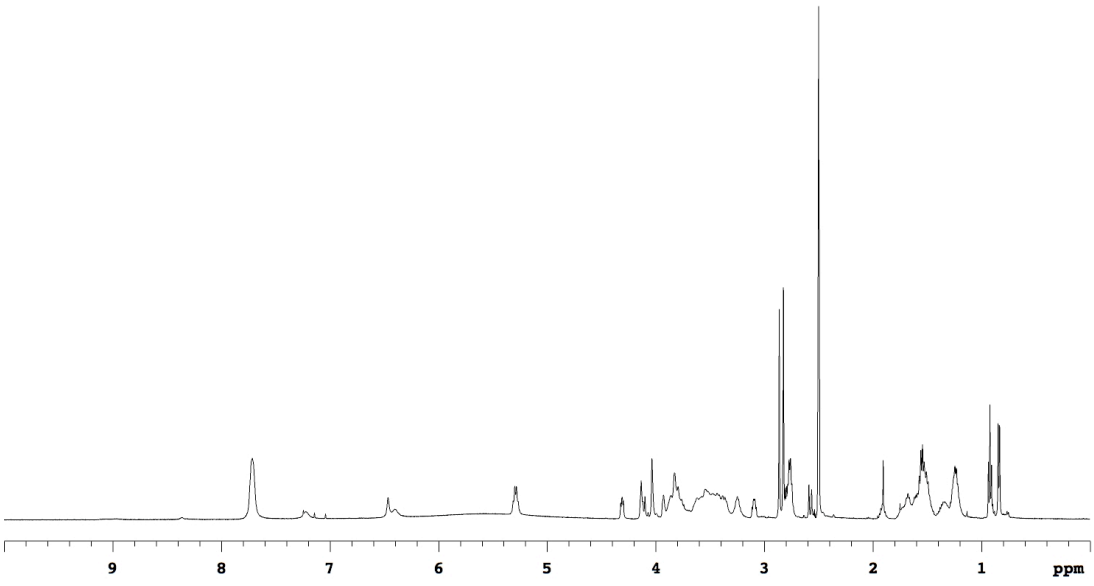
^{1}H NMR (500 MHz, DMSO-d_6) δ 7.92 (t, 2H, J = 5.6 Hz), 7.49 (br, 1H), 6.55-6.32 (m, 2H), 4.35-4.28 (m, 2H), 4.18-4.11 (m, 3H), 4.10-4.03 (m, 2H), 3.98 (d, 2H, J = 16.4 Hz), 3.88-3.81 (m, 4H), 3.81-3.58 (m, 10H), 3.52 (br, 8H), 3.32 (br, 2H), 3.13-3.07 (m, 1H), 3.07-3.01 (m, 2H), 2.96 (s, 6H), 2.82 (dd, 1H, J = 5.1, 12.4 Hz), 2.58 (d, 1H, J = 12.4 Hz), 2.33 (t, 4H, J = 7.4 Hz), 1.69-1.59 (m, 1H), 1.59-1.44 (m, 7H), 1.42-1.32 (m, 6H), 1.32-1.20 (m, 27H), 1.17 (d, 6H, J = 6.6 Hz)

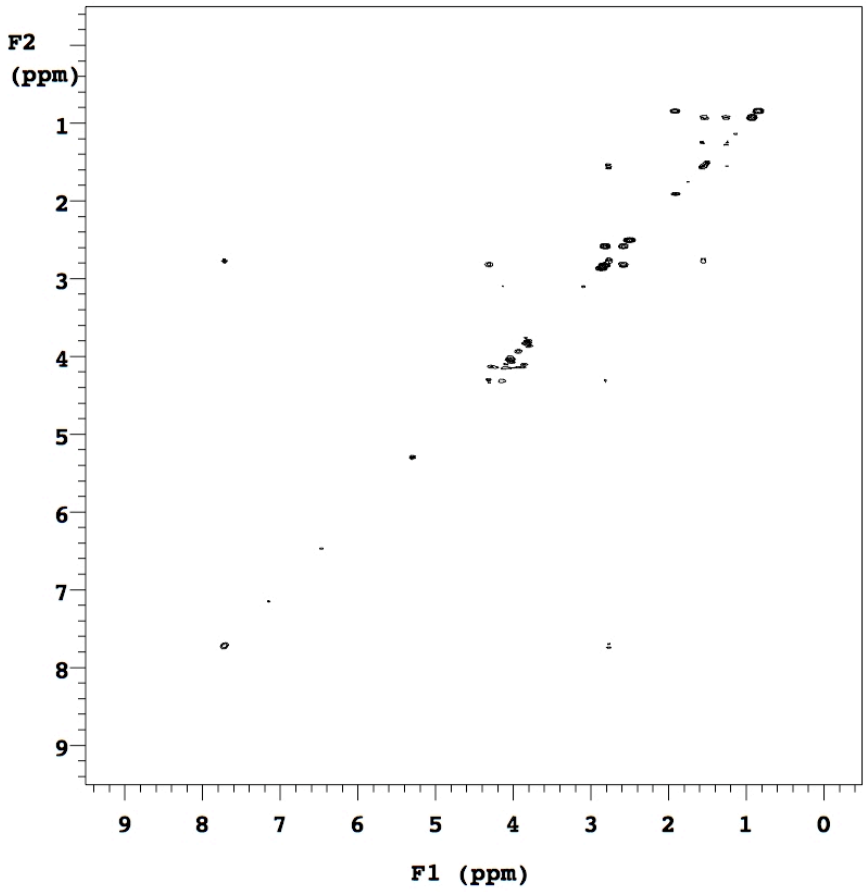

9
8
7
6
5
4
3
2
1
ppm


F2
(ppm)
1
2
3
4
5
6
7
8
9
9
8
7
6
5
4
3
2
1
0
F1 (ppm)

Data for 9ab

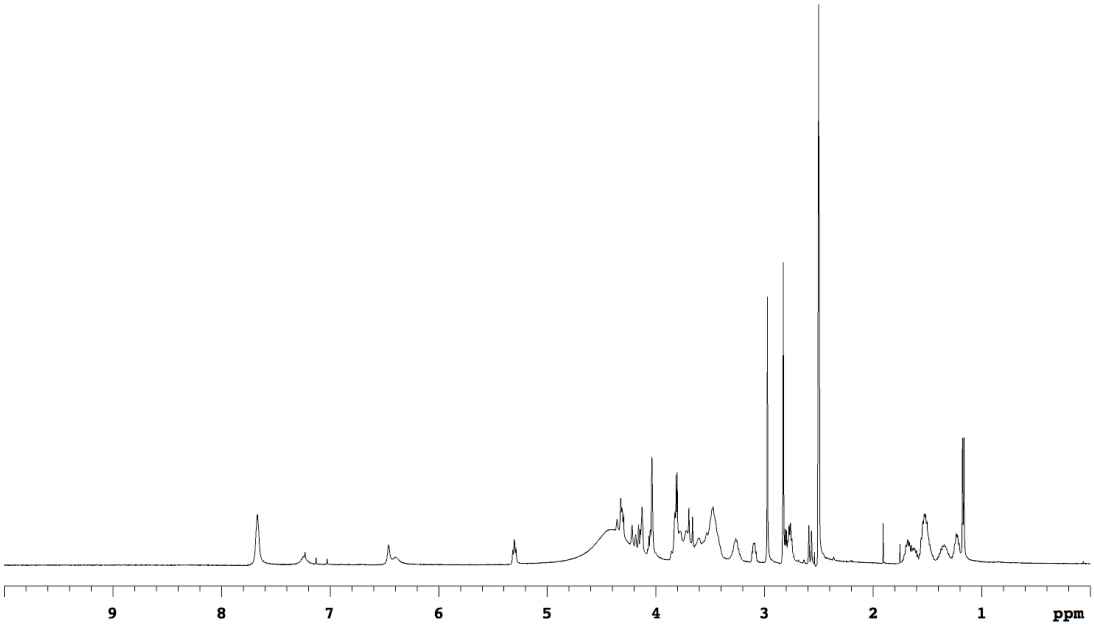
^{1}H NMR (500 MHz, DMSO-d_6) δ 7.71 (br, 4H), 7.24 (br, 1H), 6.47 (br, 1H), 6.40 (br, 1H), 5.36-5.22 (m, 2H), 4.35-4.27 (m, 1H), 4.18-4.06 (m, 2H), 4.04 (br, 2H), 3.96-3.68 (m, 8H), 3.68-3.31 (m, 12H), 3.31-3.15 (m, 2H), 3.14-3.05 (m, 1H), 2.86 (s, 3H), 2.85-2.69 (m, 8H), 2.58 (d, 1H, J = 12.5 Hz), 1.97-1.85 (m, 1H), 1.79-1.43 (m, 12H), 1.43-1.16 (m, 8H), 0.92 (t, 3H, J = 7.4 Hz), 0.84 (d, 3H, J = 6.9 Hz)

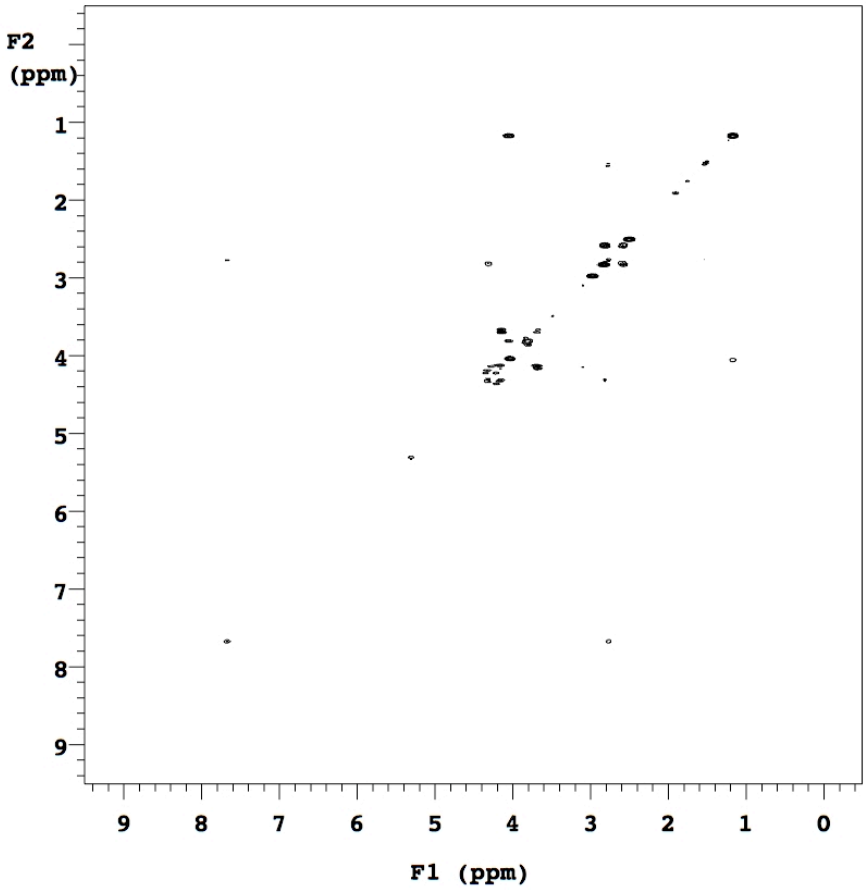




Data for 9ad

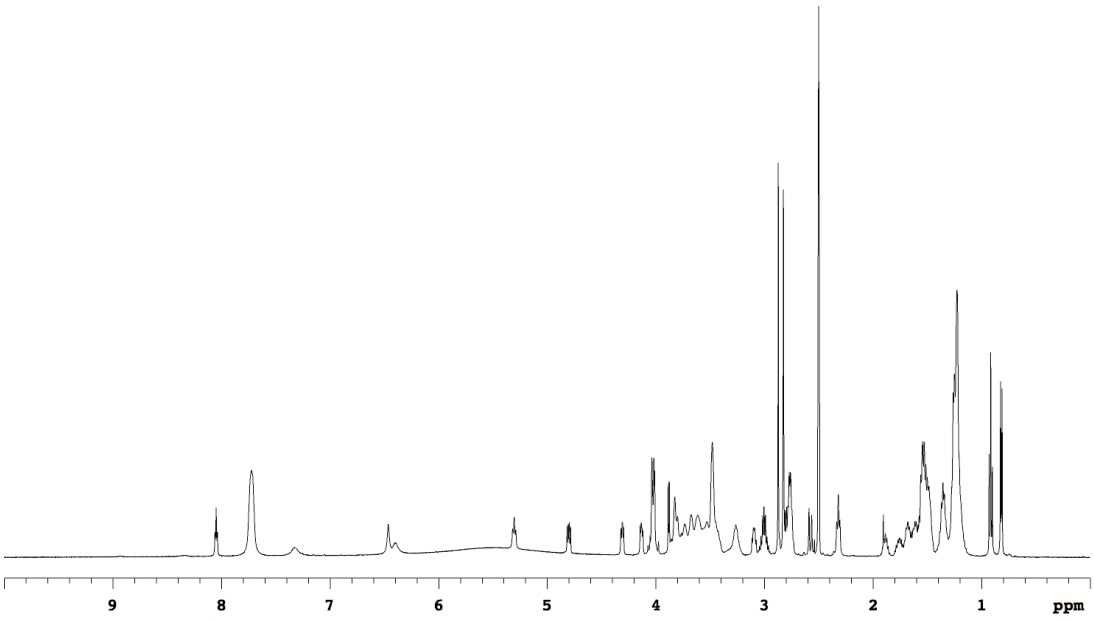
^{1}H NMR (500 MHz, DMSO-d_6) δ 7.67 (br, 2H), 7.23 (br, 1H), 6.46 (br, 1H), 6.40 (br, 1H), 5.30 (t, 1H, J = 7.7 Hz), 4.38-4.09 (m, 6H), 4.09-3.95 (m, 3H), 3.88-3.36 (m, 20H), 3.32-3.20 (m, 2H), 3.13-3.06 (m, 1H), 2.97 (s, 3H), 2.86-2.70 (m, 6H), 2.58 (d, 1H, J = 12.5 Hz), 1.75-1.59 (m, 3H), 1.59-1.43 (m, 5H), 1.43-1.28 (m, 2H), 1.28-1.19 (m, 2H), 1.17 (d, 3H, J = 6.6 Hz)

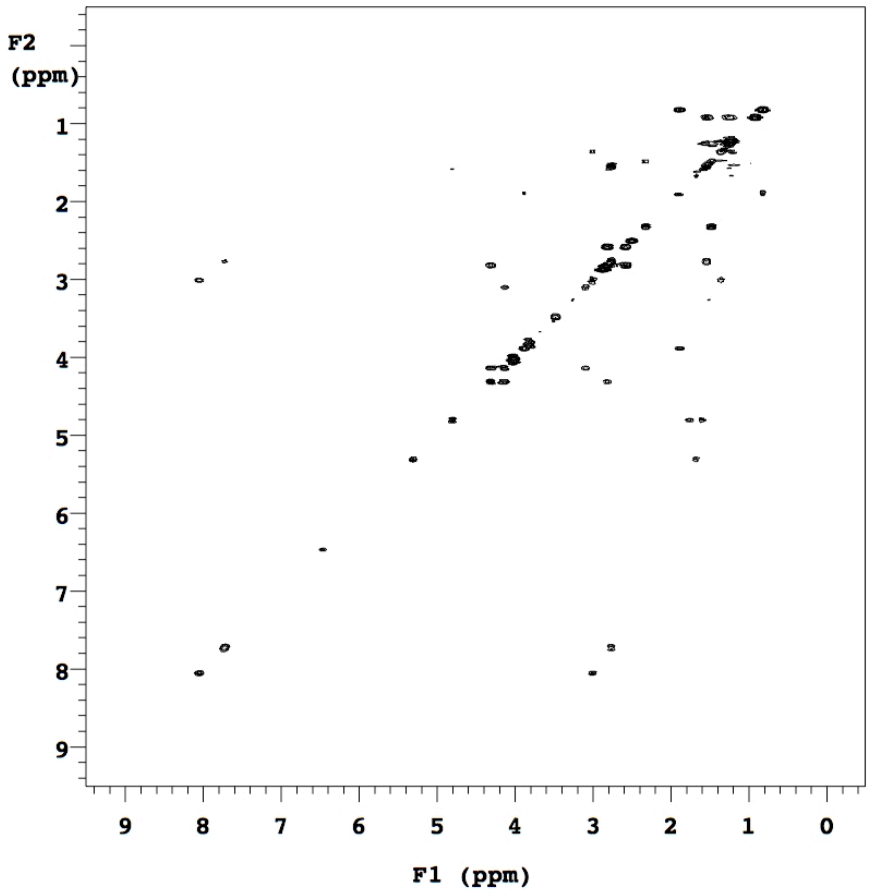

9
8
7
6
5
4
3
2
1
ppm


F2
(ppm)
1
2
3
4
5
6
7
8
9
9
8
7
6
5
4
3
2
1
0
F1 (ppm)

Data for 9af

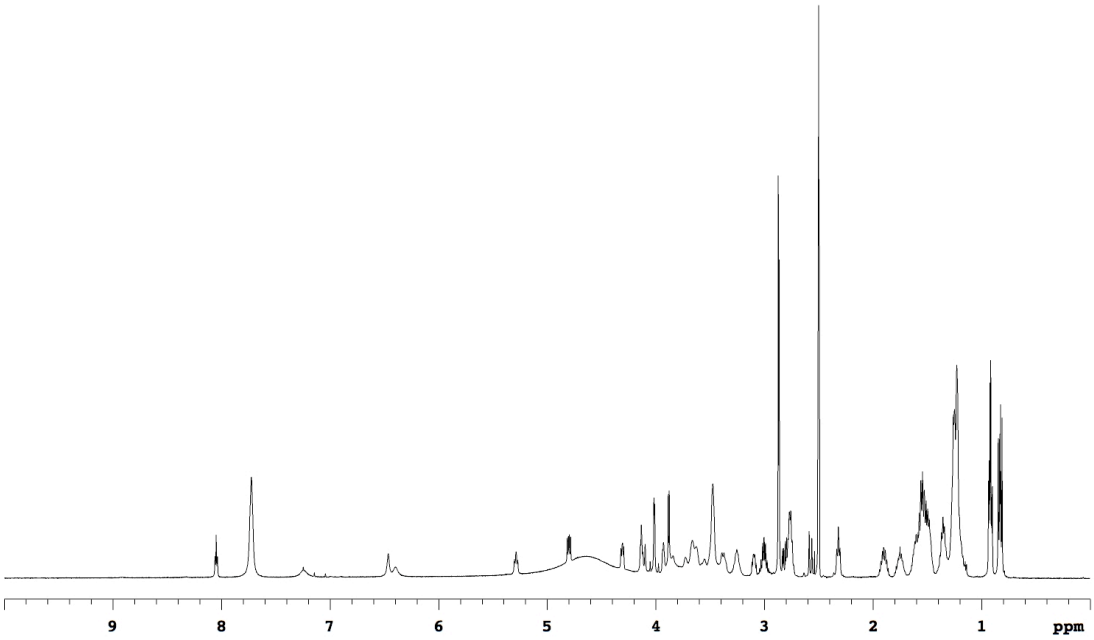
^{1}H NMR (500 MHz, DMSO-d_6) δ 8.05 (t, 1H, J =5.6 Hz), 7.72 (br, 4H), 7.33 (br, 1H), 6.46 (br, 1H), 6.40 (br, 1H), 5.30 (t, 1H, J = 7.6 Hz), 4.80 (dd, 1H, J =5.5, 10.2 Hz), 4.34-4.28 (m, 1H), 4.17-4.10 (m, 1H), 4.09-3.96 (m, 4H), 3.92-3.38 (m, 20H), 3.34-3.20 (m, 2H), 3.14-3.06 (m, 1H), 3.02-2.94 (m, 2H), 2.87 (s, 3H), 2.85-2.71 (m, 8H), 2.58 (d, 1H, J = 12.6 Hz), 2.32 (t, 2H, J = 7.2 Hz), 1.93-1.84 (m, 1H), 1.81-1.72 (m, 1H), 1.72-1.43 (m, 14H), 1.42-1.31 (m, 4H), 1.31-1.12 (m, 16H), 0.92 (t, 3H, J = 7.4 Hz), 0.82 (d, 3H, J = 6.8 Hz)

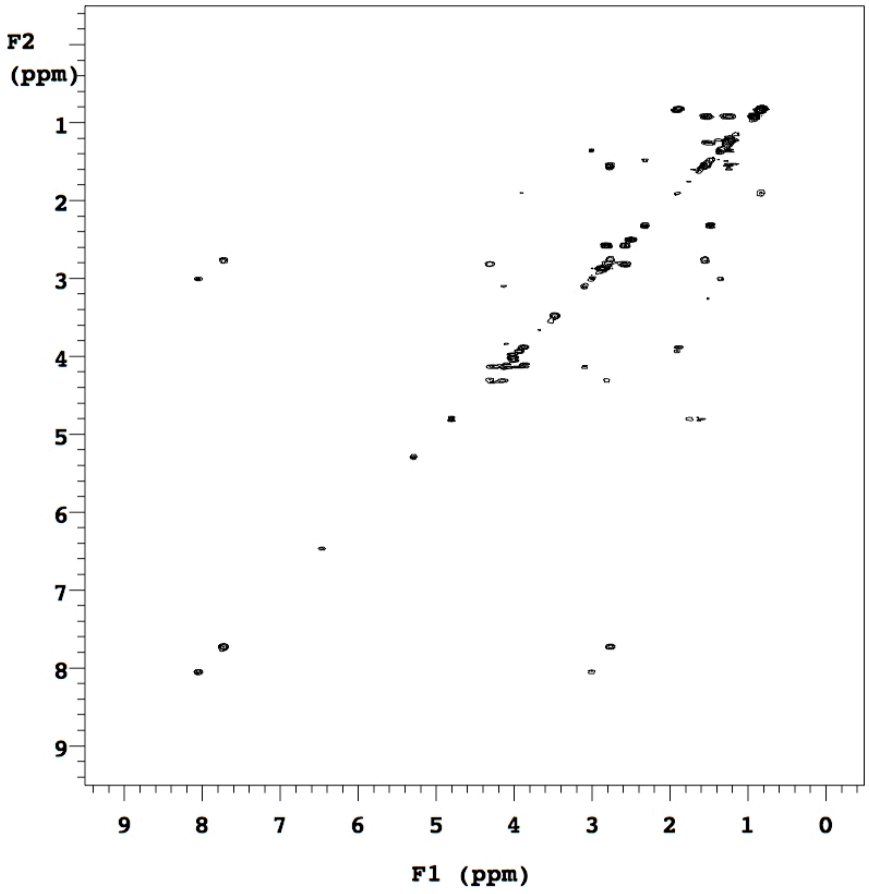




Data for 9bf

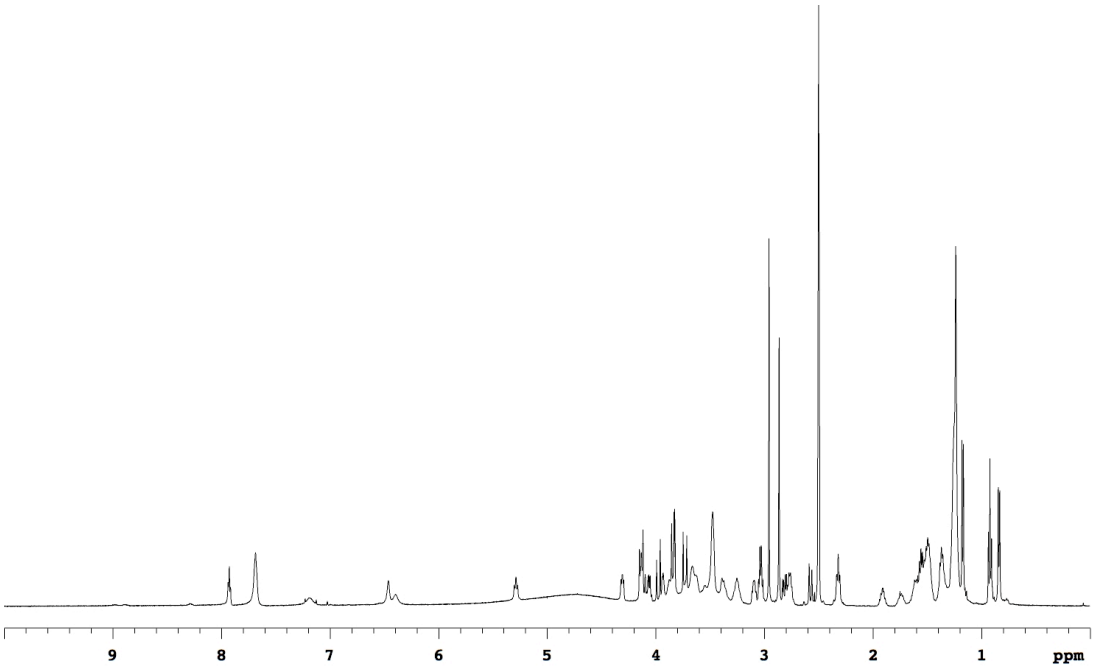
^{1}H NMR (500 MHz, DMSO-d_6) δ 8.05 (t, 1H, J = 5.5 Hz), 7.72 (br, 4H), 7.25 (br, 1H), 6.46 (br, 1H), 6.40 (br, 1H), 5.29 (t, 1H, J = 7.3 Hz), 4.80 (dd, 1H, J = 5.6, 10.3 Hz), 4.34-4.28 (m, 1H), 4.17-4.08 (m, 3H), 4.07-3.96 (m, 2H), 3.96-3.91 (m, 1H), 3.91-3.79 (m, 4H), 3.76-3.53 (m, 7H), 3.47 (br, 6H), 3.42-3.32 (m, 2H), 3.31-3.19 (m, 2H), 3.13-3.06 (m, 1H), 3.06-2.95 (m, 2H), 2.90-2.85 (m, 6H), 2.85-2.71 (m, 5H), 2.58 (d, 1H, J = 12.4 Hz), 2.32 (t, 2H, J = 7.2 Hz), 1.96-1.85 (m, 1H), 1.80-1.69 (m, 1H), 1.67-1.43 (m, 14H), 1.42-1.31 (m, 4H), 1.30-1.13 (m, 18H), 0.96-0.89 (m, 6H), 0.87-0.79 (m, 6H)

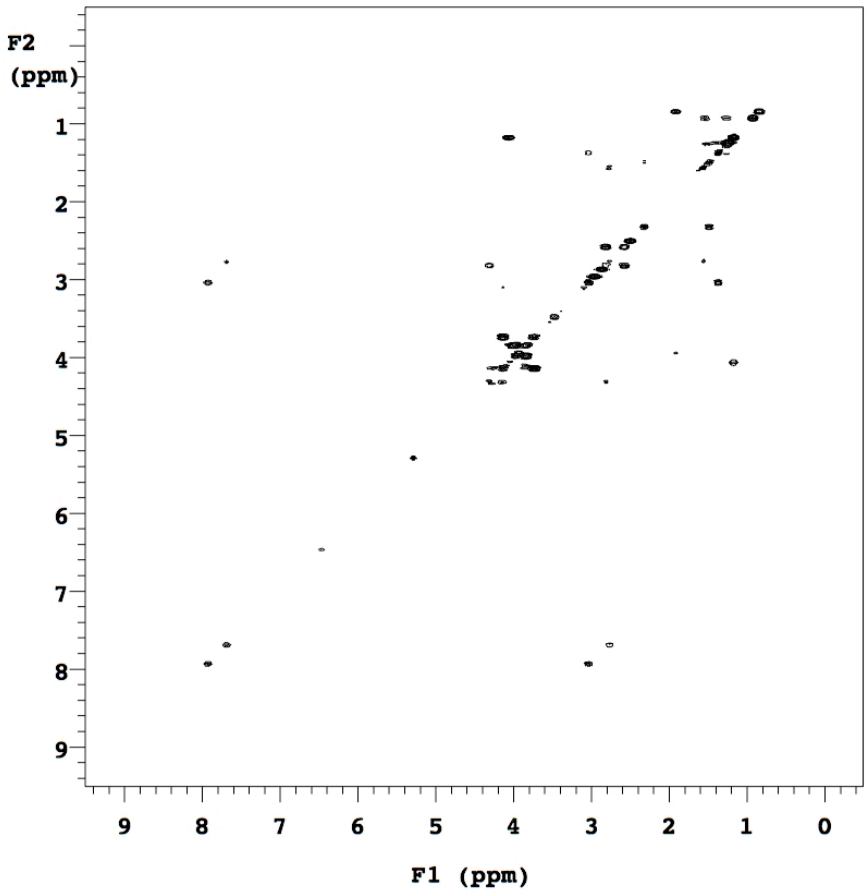

9
8
7
6
5
4
3
2
1
ppm


F2
(ppm)
1
2
3
4
5
6
7
8
9
9
8
7
6
5
4
3
2
1
0
F1 (ppm)

Data for 9bh

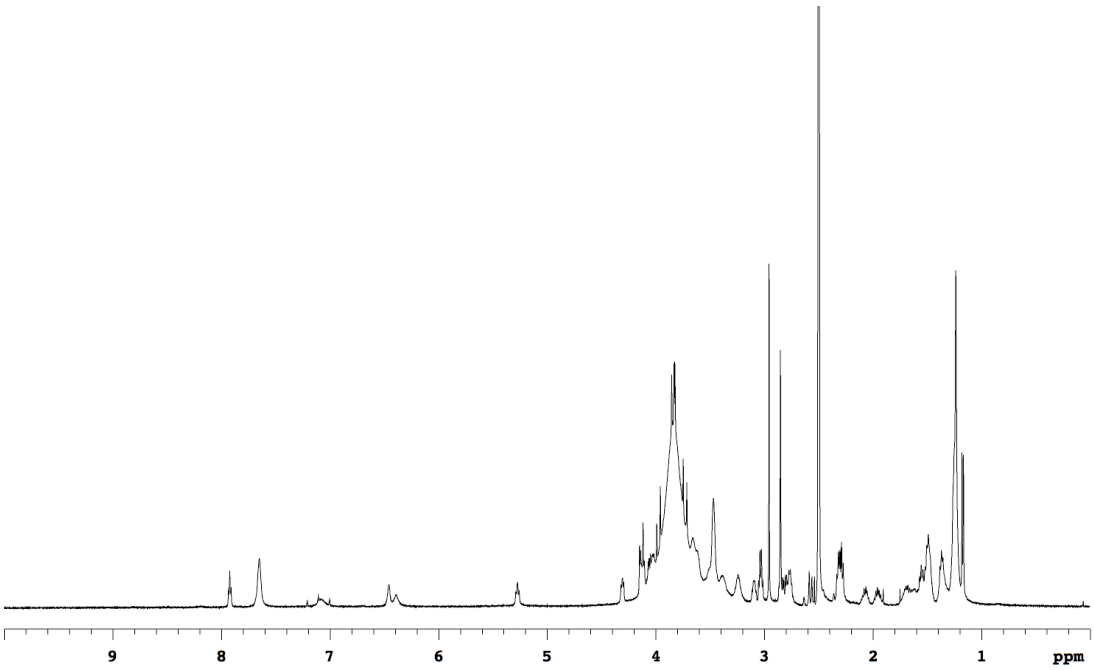
^{1}H NMR (500 MHz, DMSO-d_6) δ 7.93 (t, 1H, J = 5.5 Hz), 7.68 (br, 2H), 7.18 (br, 1H), 6.46 (br, 1H), 6.39 (br, 1H), 5.29 (t, 1H, J = 7.3 Hz), 4.35-4.26 (m, 1H), 4.19-4.02 (m, 4H), 4.02-3.91 (m, 3H), 3.91-3.78 (m, 4H), 3.78-3.59 (m, 7H), 3.59-3.43 (m, 8H), 3.43-3.31 (m, 2H), 3.31-3.16 (m, 2H), 3.14-3.07 (m, 1H), 3.07-2.99 (m, 2H), 2.96 (s, 3H), 2.86 (s, 3H), 2.84-2.69 (m, 3H), 2.58 (d, 1H, J = 12.5 Hz), 2.32 (t, 2H, J = 7.2 Hz), 1.96-1.87 (m, 1H), 1.80-1.69 (m, 1H), 1.67-1.43 (m, 10H), 1.43-1.31 (m, 4H), 1.30-1.20 (m, 14H), 1.17 (d, 3H, J = 6.5 Hz), 0.92 (t, 3H, J = 7.4 Hz), 0.84 (d, 3H, J = 6.8 Hz)

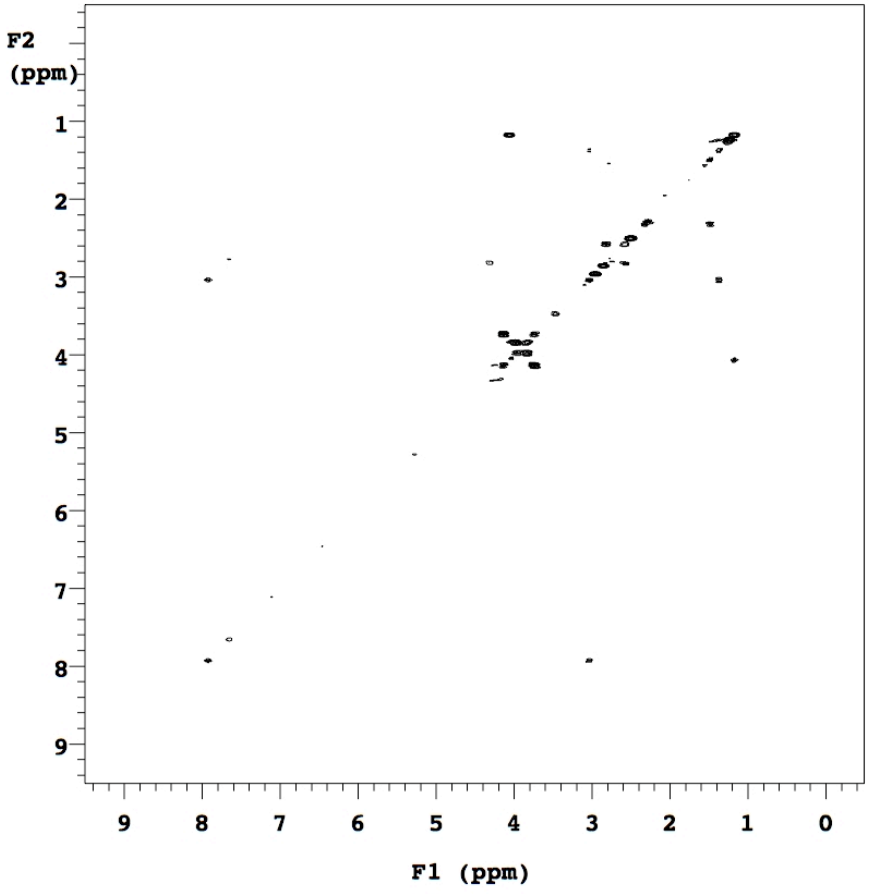




Data for 9ch

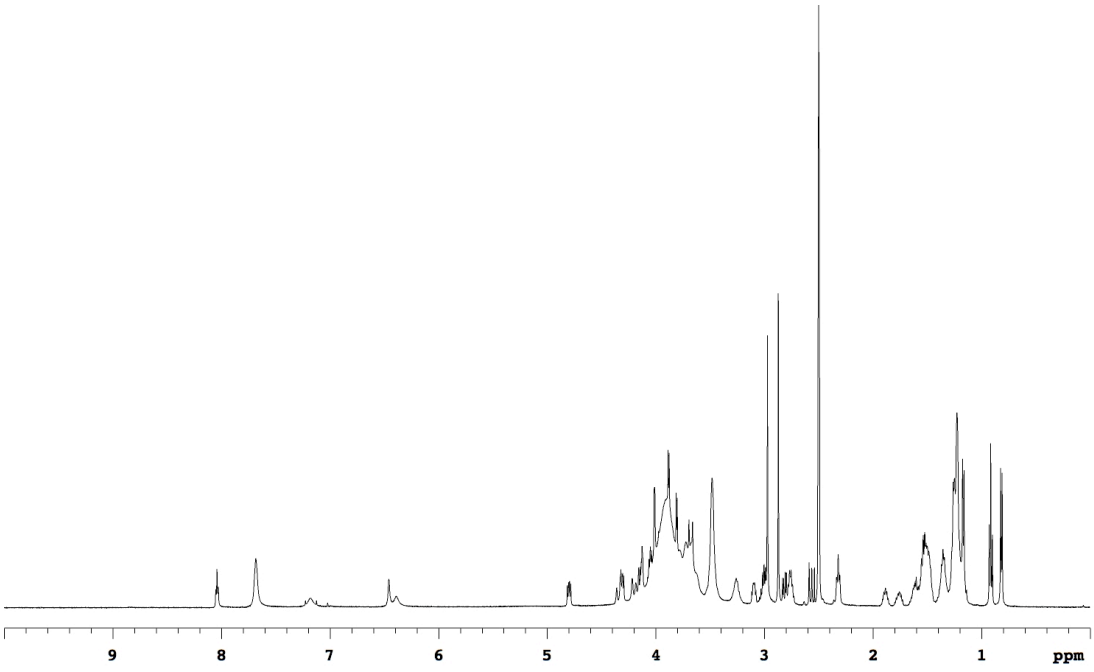
^{1}H NMR (500 MHz, DMSO-d_6) δ 7.92 (t, 1H, J = 5.5 Hz), 7.65 (br, 2H), 7.11 (br, 1H), 6.46 (br, 1H), 6.39 (br, 1H), 5.27 (t, 1H, J = 7.8 Hz), 4.36-4.27 (m, 1H), 4.18-4.09 (m, 2H), 4.09-3.56 (m, 17H), 3.54-3.31 (m, 10H), 3.24 (br, 2H), 3.13-3.07 (m, 1H), 3.06-3.00 (m, 2H), 2.96 (s, 3H), 2.88-2.72 (m, 6H), 2.57 (d, 1H, J = 12.5 Hz), 2.35-2.25 (m, 4H), 2.13-2.02 (m, 1H), 2.01-1.89 (m, 1H), 1.76-1.43 (m, 9H), 1.41-1.31 (m, 3H), 1.30-1.20 (m, 11H), 1.17 (d, 6H, J = 6.6 Hz)

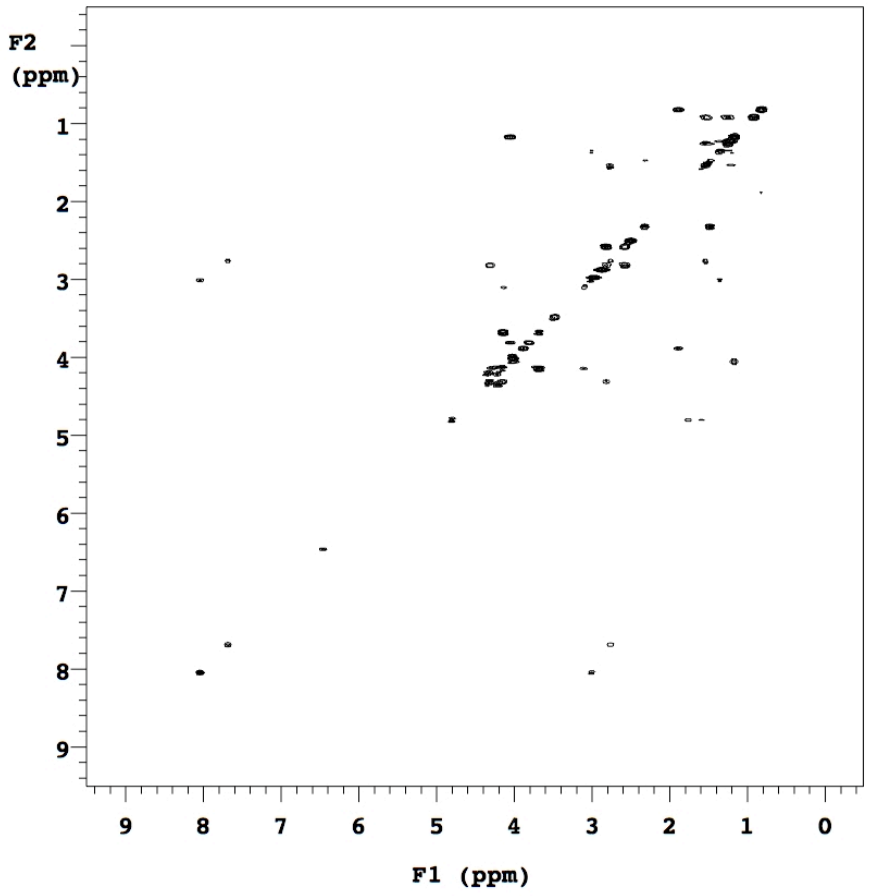

9
8
7
6
5
4
3
2
1
ppm


F2
(ppm)
1
2
3
4
5
6
7
8
9
9
8
7
6
5
4
3
2
1
0
F1 (ppm)

Data for 9df

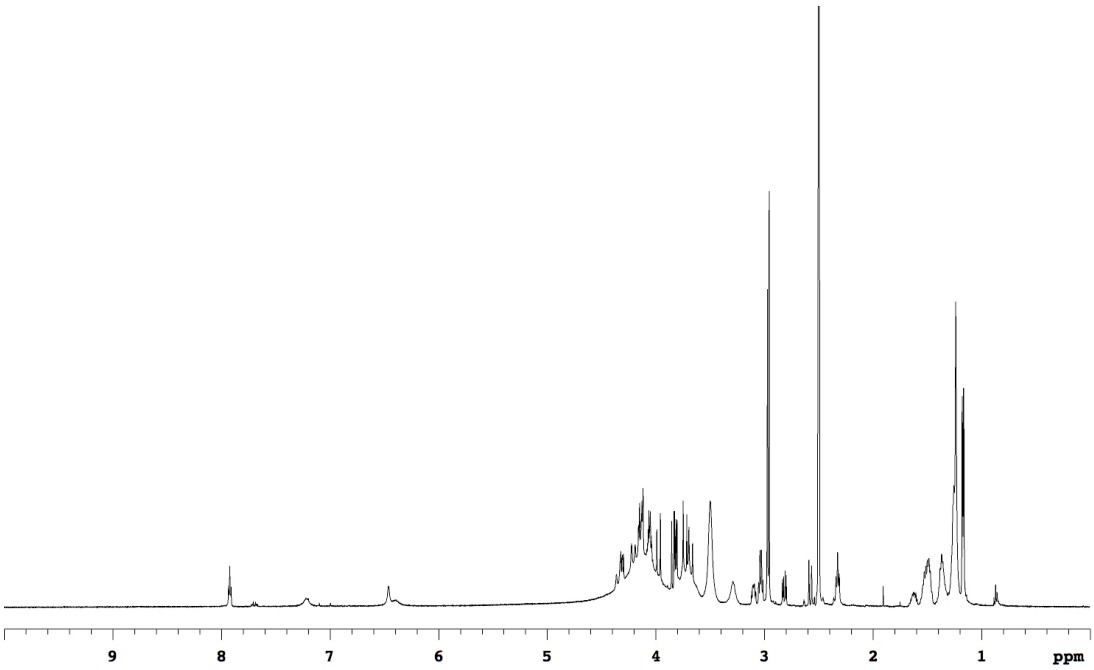
^{1}H NMR (500 MHz, DMSO-d_6) δ 8.04 (t, 1H, J = 5.5 Hz), 7.68 (br, 2H), 7.19 (br, 1H), 6.46 (br, 1H), 6.39 (br, 1H), 4.80 (dd, 1H, J = 5.7, 10.1 Hz), 4.38-4.27 (m, 2H), 4.25-4.10 (m, 4H), 4.09-3.57 (m, 16H), 3.48 (br, 8H), 3.32-3.19 (m, 2H), 3.13-3.06 (m, 1H), 3.06-2.93 (m, 5H), 2.87 (s, 3H), 2.85-2.71 (m, 3H), 2.58 (d, 1H, J = 12.5 Hz), 2.32 (t, 2H, J = 7.2 Hz), 1.93-1.84 (m, 1H), 1.81-1.71 (m, 1H), 1.68-1.44 (m, 9H), 1.41-1.31 (m, 4H), 1.31-1.11 (m, 17H), 0.92 (t, 3H, J = 7.4 Hz), 0.82 (d, 3H, J = 6.8 Hz)

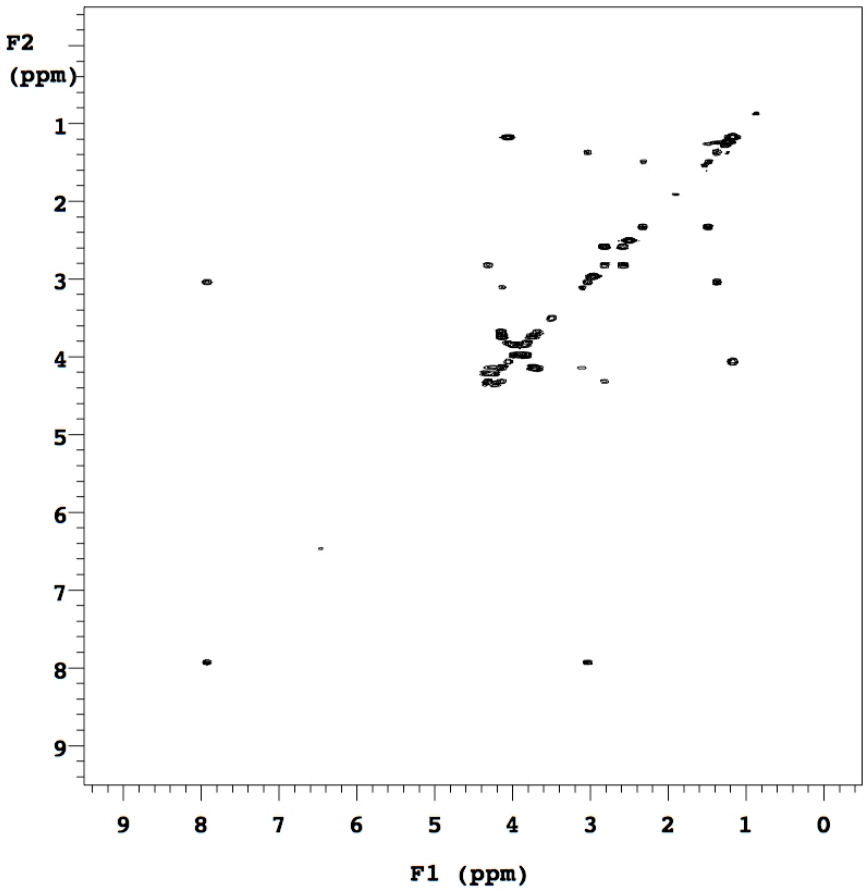




Data for 9dh

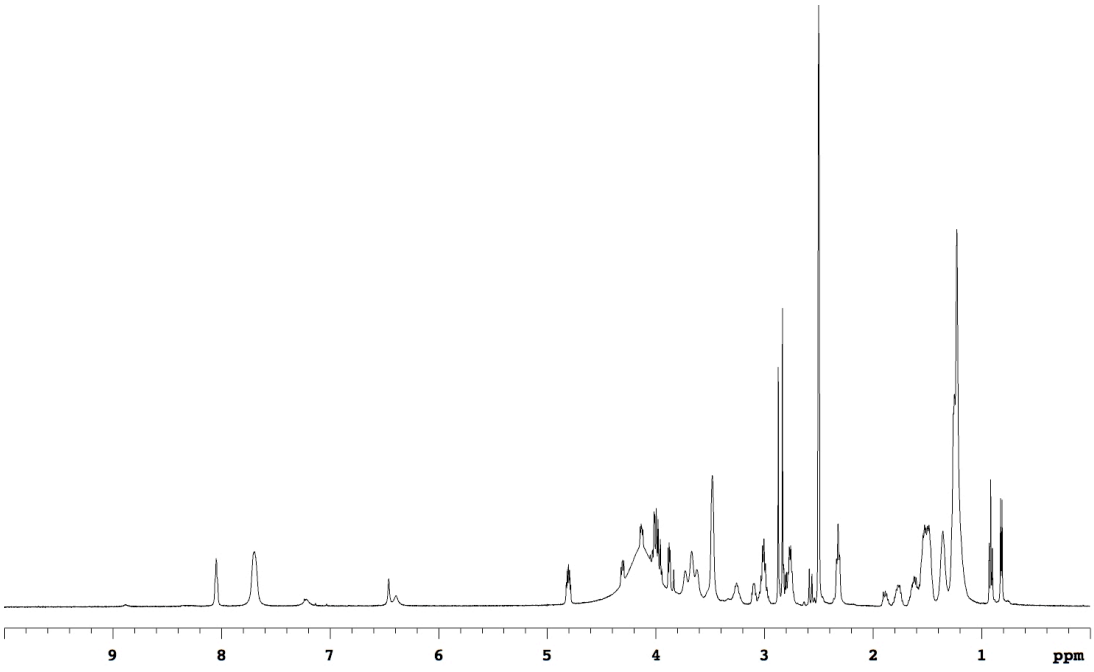
^{1}H NMR (500 MHz, DMSO-d_6) δ 7.92 (t, 1H, J = 5.5 Hz), 7.22 (br, 1H), 6.46 (br, 1H), 6.40 (br, 1H), 4.39-4.27 (m, 3H), 4.25-3.94 (m, 7H), 4.02-3.91 (m, 3H), 3.88-3.58 (m, 12H), 3.50 (br, 8H), 3.29 (br, 2H), 3.13-3.07 (m, 1H), 3.07-3.00 (m, 2H), 2.97 (s, 3H), 2.96 (s, 3H), 2.82 (dd, 1H, J = 5.2, 12.5 Hz), 2.58 (d, 1H, J = 12.5 Hz), 2.33 (t, 2H, J = 7.4 Hz), 1.68-1.58 (m, 1H), 1.57-1.44 (m, 4H), 1.42-1.31 (m, 4H), 1.30-1.20 (m, 12H), 1.19-1.15 (m, 6H)

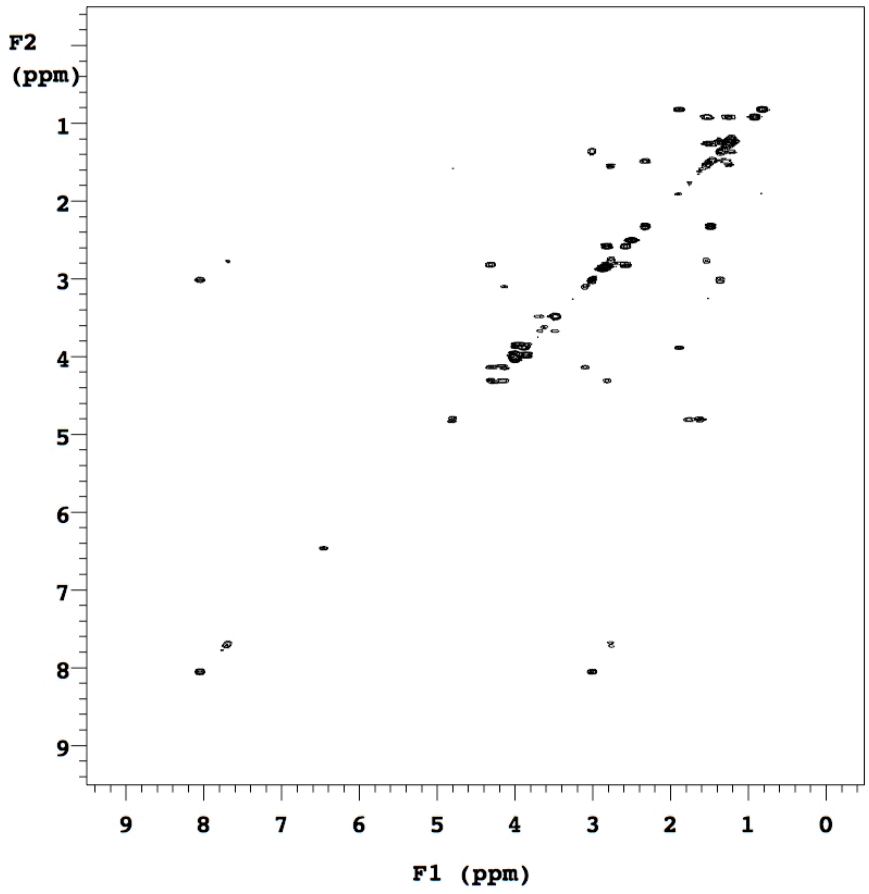

9
8
7
6
5
4
3
2
1
ppm


F2
(ppm)
1
2
3
4
5
6
7
8
9
9
8
7
6
5
4
3
2
1
0
F1 (ppm)

Data for 9ef

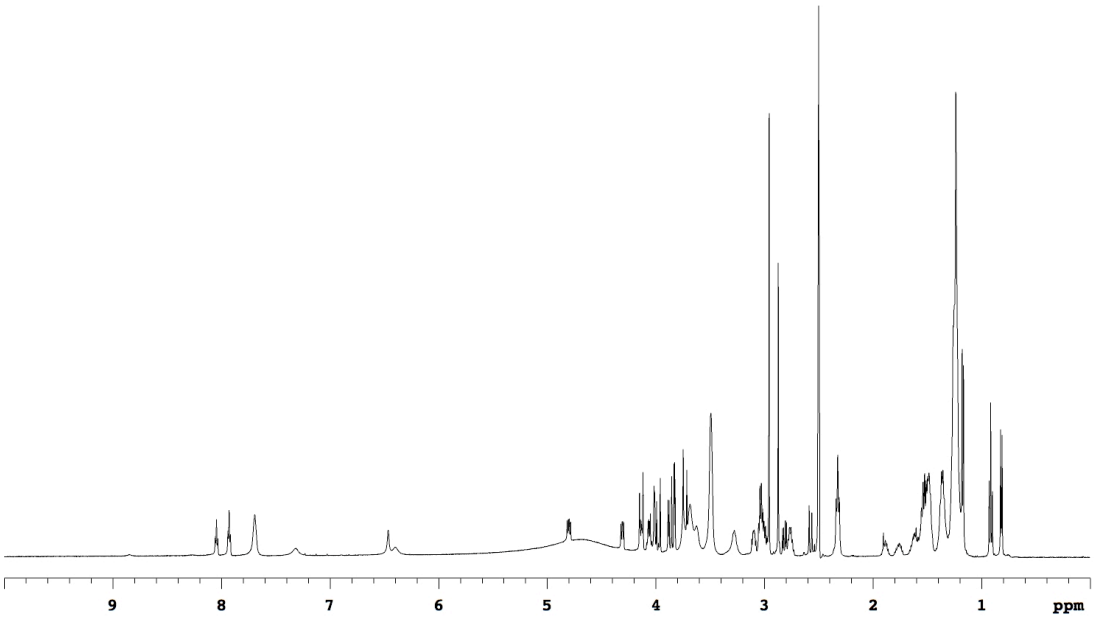
^{1}H NMR (500 MHz, DMSO-d_6) δ 8.05 (br, 2H), 7.70 (br, 4H), 7.24 (br, 1H), 6.46 (br, 1H), 6.39 (br, 1H), 4.86-4.76 (m, 2H), 4.34-4.28 (m, 1H), 4.16-3.93 (m, 5H), 3.91-3.82 (m, 3H), 3.78-3.56 (m, 8H), 3.48 (br, 8H), 3.32-3.19 (m, 2H), 3.14-3.07 (m, 1H), 3.07-2.95 (m, 4H), 2.87 (s, 3H), 2.85-2.70 (m, 8H), 2.57 (d, 1H, J = 12.5 Hz), 2.32 (t, 4H, J = 7.2 Hz), 1.93-1.84 (m, 1H), 1.83-1.71 (m, 2H), 1.69-1.43 (m, 14H), 1.41-1.31 (m, 6H), 1.31-1.09 (m, 30H), 0.92 (t, 3H, J = 7.4 Hz), 0.82 (d, 3H, J = 6.8 Hz)

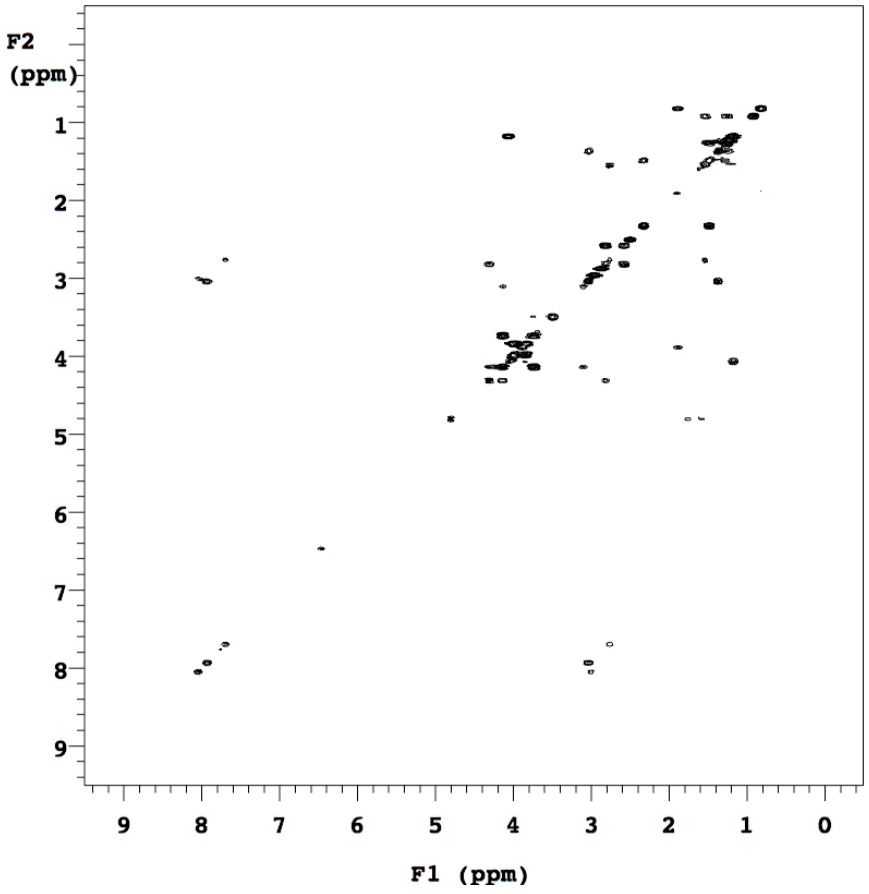




Data for 9fh

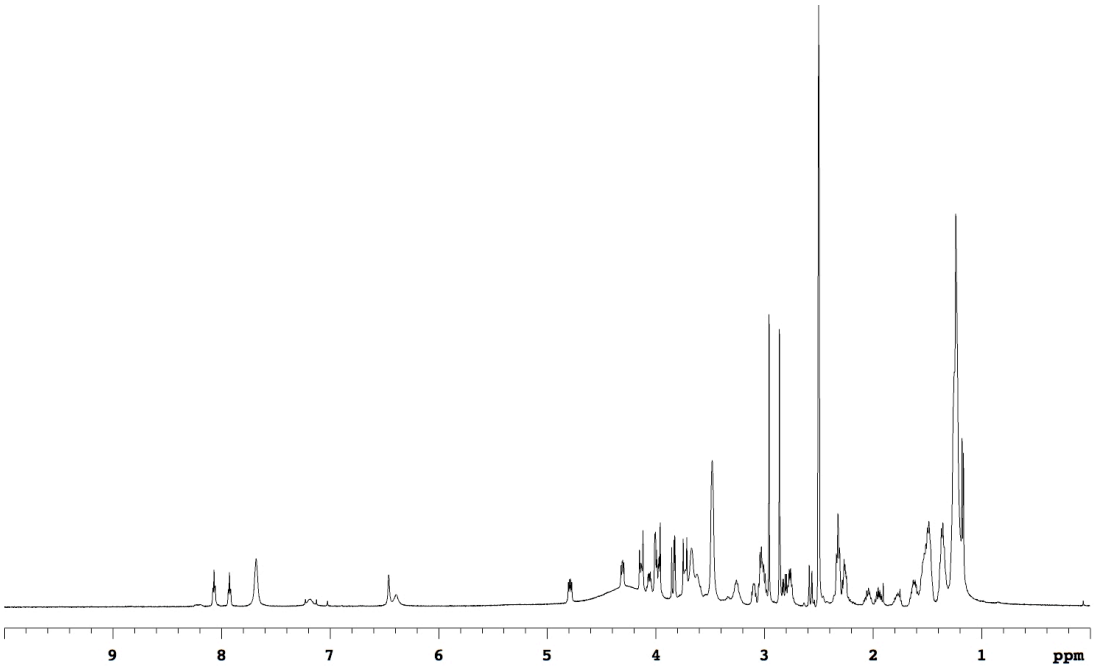
^{1}H NMR (500 MHz, DMSO-d_6) δ 8.04 (t, 1H, J = 5.6 Hz), 7.93 (t, 1H, J = 5.6 Hz), 7.69 (br, 2H), 7.32 (br, 1H), 6.46 (br, 1H), 6.40 (br, 1H), 4.80 (dd, 1H, J = 5.5, 10.0 Hz), 4.34-4.28 (m, 1H), 4.17-4.10 (m, 3H), 4.09-4.03 (m, 2H), 4.03-3.94 (m, 3H), 3.91-3.80 (m, 4H), 3.80-3.58 (m, 7H), 3.49 (br, 8H), 3.35-3.21 (m, 2H), 3.13-3.07 (m, 1H), 3.07-2.97 (m, 4H), 2.96 (s, 3H), 2.87 (s, 3H), 2.85-2.72 (m, 3H), 2.58 (d, 1H, J = 12.5 Hz), 2.32 (t, 4H, J = 7.2 Hz), 1.93-1.84 (m, 1H), 1.81-1.72 (m, 1H), 1.68-1.43 (m, 11H), 1.42-1.31 (m, 6H), 1.31-1.12 (m, 31H), 0.92 (t, 3H, J = 7.4 Hz), 0.82 (d, 3H, J = 6.8 Hz)

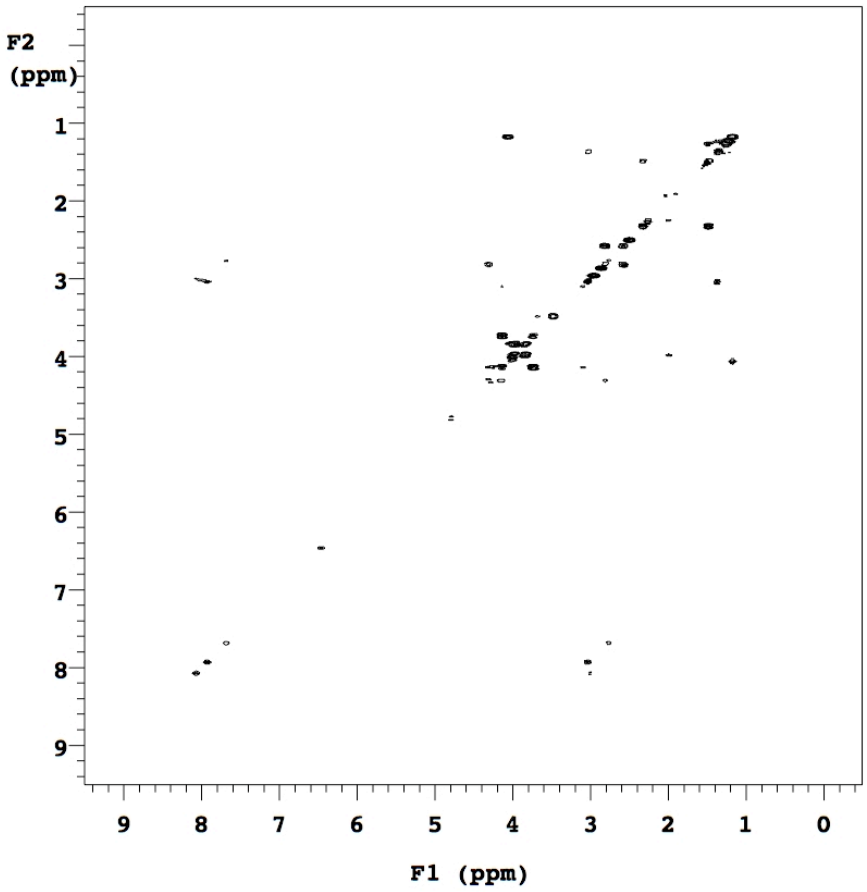

9
8
7
6
5
4
3
2
1
ppm


F2
(ppm)
1
2
3
4
5
6
7
8
9
9
8
7
6
5
4
3
2
1
0
F1 (ppm)

Data for 9gh

^{1}H NMR (500 MHz, DMSO-d_6) δ 8.07 (t, 1H, J = 5.6 Hz), 7.93 (t, 1H, J = 5.6 Hz), 7.68 (br, 2H), 7.18 (br, 1H), 6.46 (br, 1H), 6.39 (br, 1H), 4.79 (dd, 1H, J = 5.4, 10.1 Hz), 4.34-4.28 (m, 1H), 4.17-4.10 (m, 2H), 4.09-4.03 (m, 2H), 4.03-3.92 (m, 4H), 3.87-3.81 (m, 3H), 3.78-3.57 (m, 9H), 3.48 (br, 7H), 3.30-3.19 (m, 2H), 3.13-3.07 (m, 1H), 3.07-3.97 (m, 4H), 2.96 (s, 3H), 2.86 (s, 3H), 2.84-2.72 (m, 3H), 2.57 (d, 1H, J = 12.5 Hz), 2.38-2.22 (m, 6H), 2.10-2.00 (m, 1H), 1.99-1.89 (m, 1H), 1.83-1.73 (m, 1H), 1.68-1.58 (m, 2H), 1.58-1.43 (m, 9H), 1.41-1.31 (m, 6H), 1.31-1.10 (m, 29H)

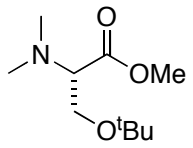



F2
(ppm)
1
2
3
4
5
6
7
8
9
9
8
7
6
5
4
3
2
1
0
F1 (ppm)

APPENDIX C

EXPERIMENTAL FOR CHAPTER IV

General Methods. All chemicals were obtained from commercial suppliers and used without further purification. Dichloromethane was obtained anhydrous by distillation over calcium hydride and THF was distilled over sodium metal and benzophenone. Analytical HPLC analyses were carried out on 25 x 0.46 cm C-18 column using gradient conditions (5 – 95% B). The eluents used were: solvent A (H_2O with 0.1% TFA) and solvent B (CH_3CN with 0.1% TFA). The flow rate used was 1.0 mL/min. NMR spectra were recorded at 300 MHz and 500 MHz. NMR chemical shifts were expressed in ppm relative to internal solvent peaks, and coupling constants were measured in Hz.

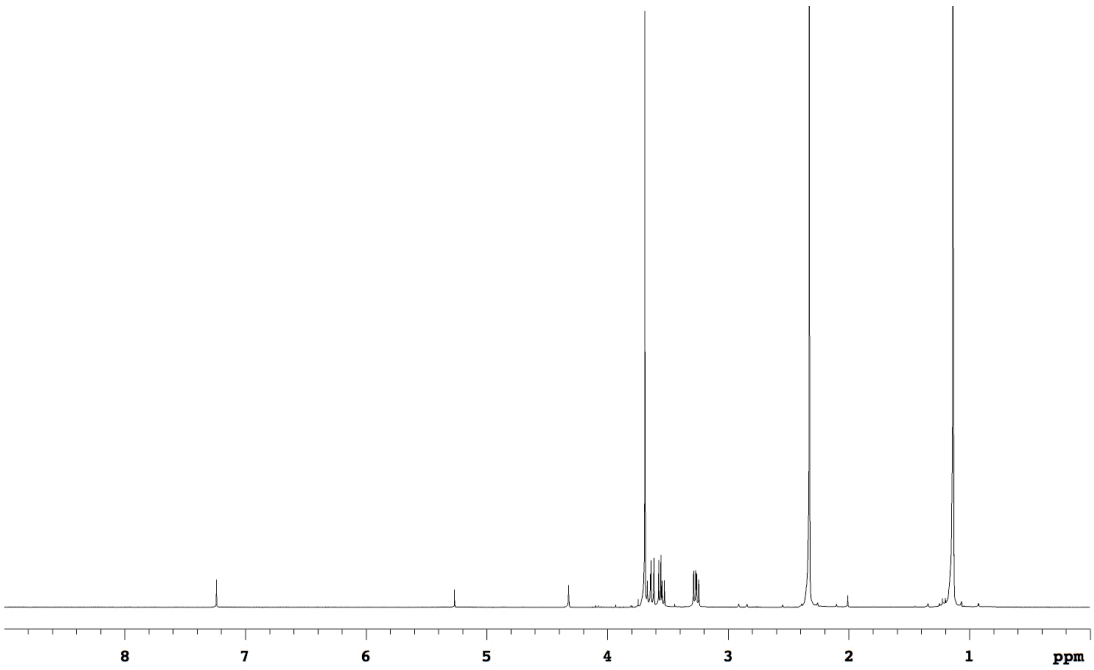


13

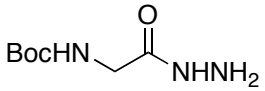
Preparation of Compound 13. To a solution of 1.06 g (5.0 mmol) H-Ser(*t*Bu)-OMe in 16 mL CH_3OH was added acetic acid to adjust the pH (~4.0). To this solution was slowly added 4 mL (50.0 mmol) formaldehyde solution (concentration, 37 %), and then 471.3 mg (7.5 mmol) $NaCNBH_3$ under stirring. The reaction mixture was stirred under room temperature for 10 h. The solvents were removed under reduced pressure. The residue was washed with sat. $NaHCO_3$ and extracted with CH_2Cl_2. The CH_2Cl_2 solution was dried with Na_2SO_4 and concentrated to dryness. Flash chromatography with EtOAc/hexanes mixtures (20 % to 30 %) afforded 761.3 mg (75 %) pure *N*, *N*-demethyl-Ser(*t*Bu)-OMe as a colorless liquid.

^{1}H NMR (300 MHz, $CDCl_3$) δ 3.69 (s, 3H), 3.64 (dd, 1H, J = 7.4, 8.8 Hz), 3.55 (dd, 1H, J = 5.6, 8.8 Hz), 3.26 (dd, 1H, J = 5.6, 7.4 Hz), 2.33 (s, 6H), 1.14 (s, 9H)

MS (APCI, m/z) 204.0 $(M+H)^+$

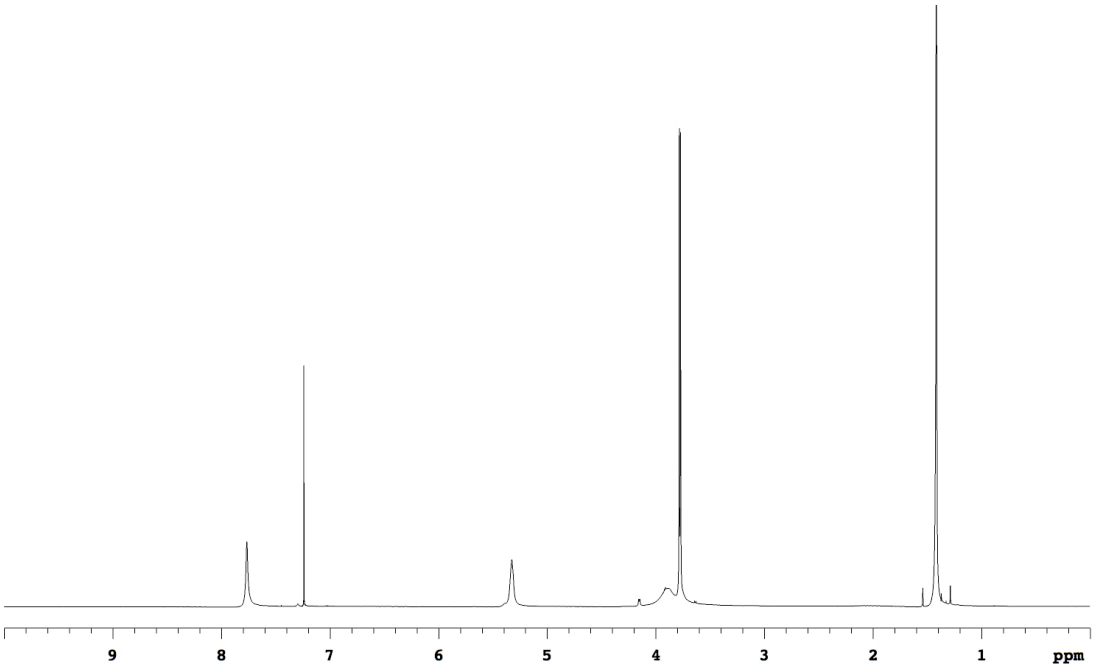


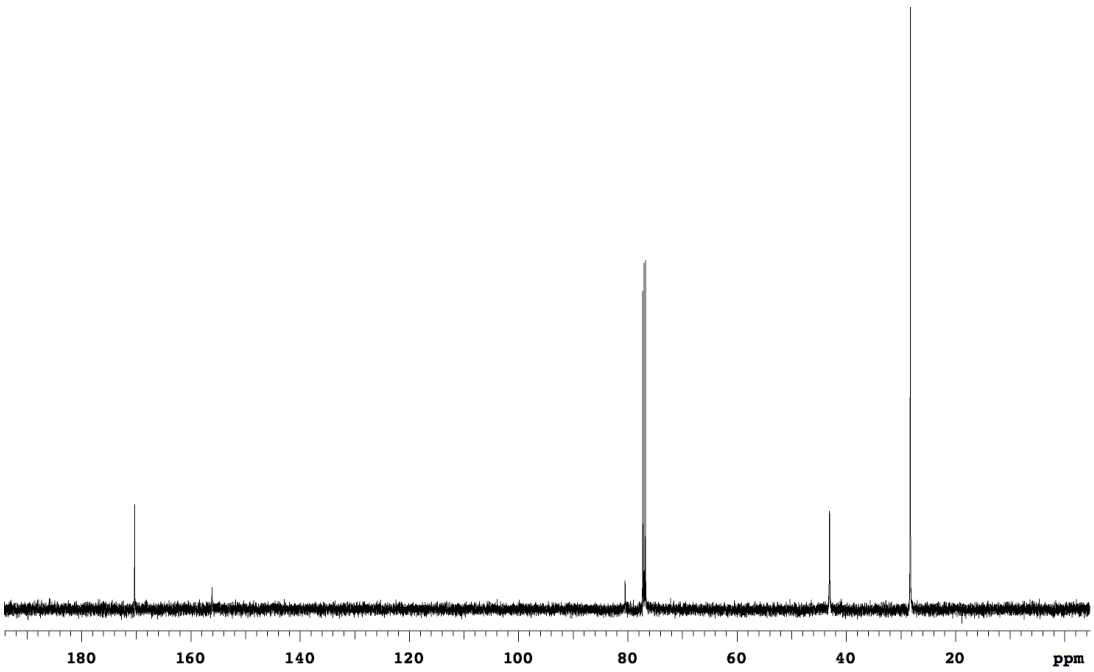
General Procedure for Preparation of Compound 10a-f. A solution of *N*-Boc protected L-amino acid methyl ester (1.0 equiv) in ethanol (0.4 M) was added hydrazine monohydrate (3 equiv) in one portion. The reaction mixture was vigorously stirred at 25 °C for 12 h. The solvents were evaporated *in vacuo* and the crude amino acid hydrazide product was used in the next step without further purification.

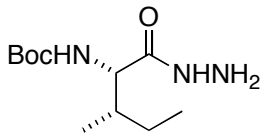


10a

Compound **10a** was prepared from Boc-Gly-OMe (3.78 g, 20.0 mmol) as a white solid.
^{1}H NMR (300 MHz, $CDCl_3$) δ 7.77 (br, 1H), 5.33 (br, 1H), 3.89 (br, 2H), 3.78 (d, 2H, J = 6.0 Hz), 1.42 (s, 9H)
^{13}C NMR (125 MHz, $CDCl_3$) δ 170.3, 156.1, 80.5, 43.1, 28.3


9
8
7
6
5
4
3
2
1
ppm


180
160
140
120
100
80
60
40
20
ppm

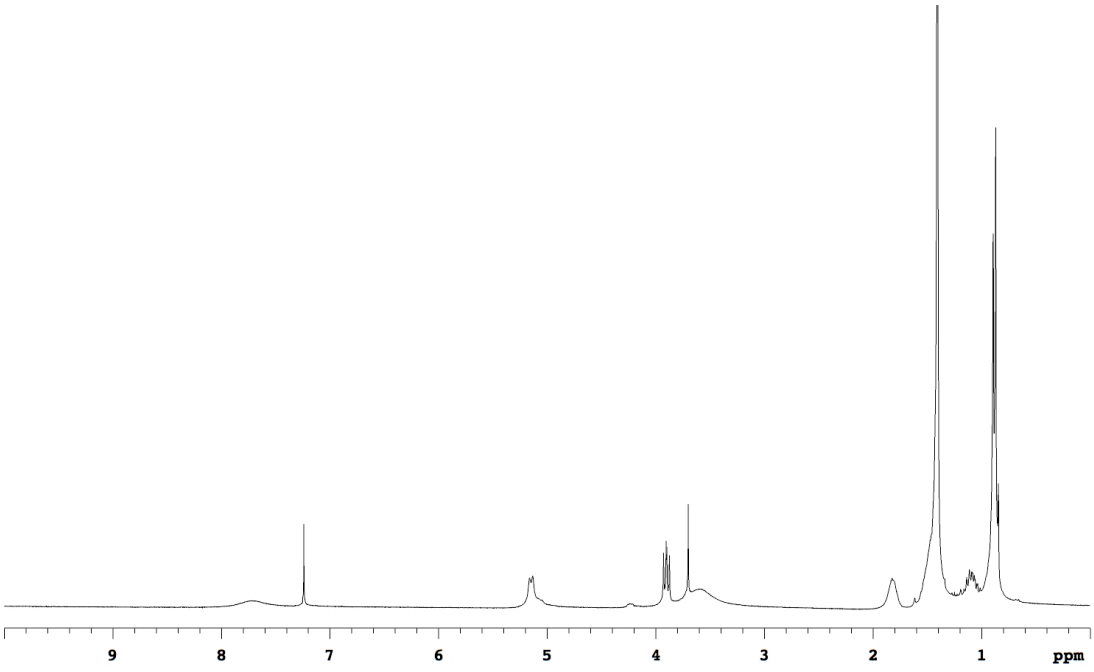


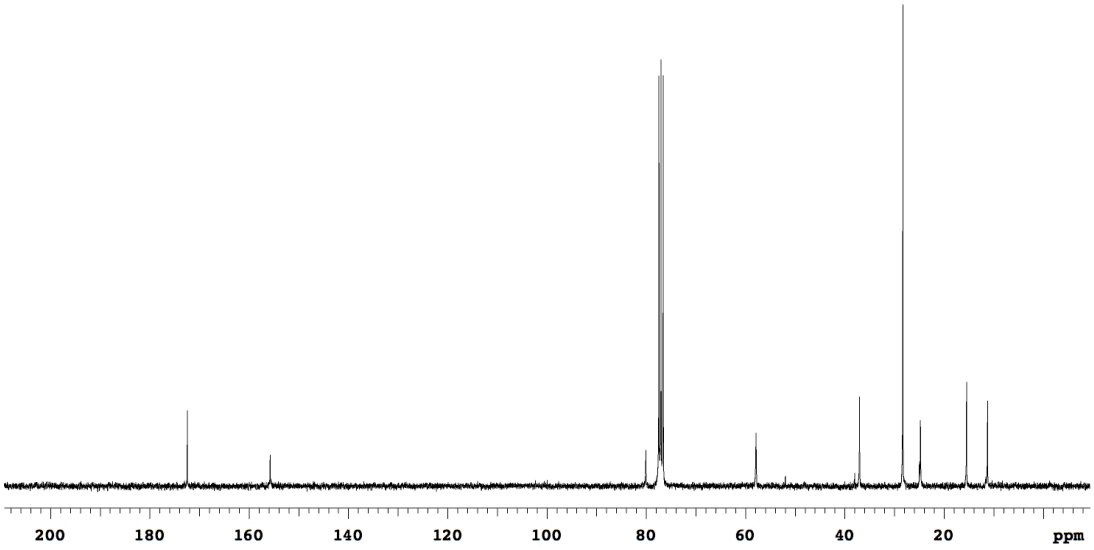
10b

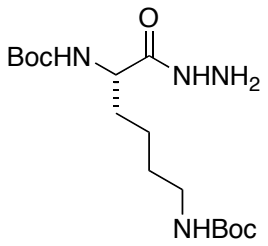
Compound **10b** was prepared from Boc-Ile-OMe (3.68 g, 15.0 mmol) as a white solid.

^{1}H NMR (300 MHz, $CDCl_3$) δ 7.70 (br, 1H), 5.15 (d, 1H, J = 8.6 Hz), 4.02-3.30 (m, 3H), 1.82 (br, 1H), 1.60-1.27 (m, 10H), 1.18-1.01 (m, 1H), 1.00-0.75 (m, 6H)

^{13}C NMR (75 MHz, $CDCl_3$) δ 172.4, 155.8, 80.1, 57.9, 37.0, 28.3, 24.8, 15.5, 11.2





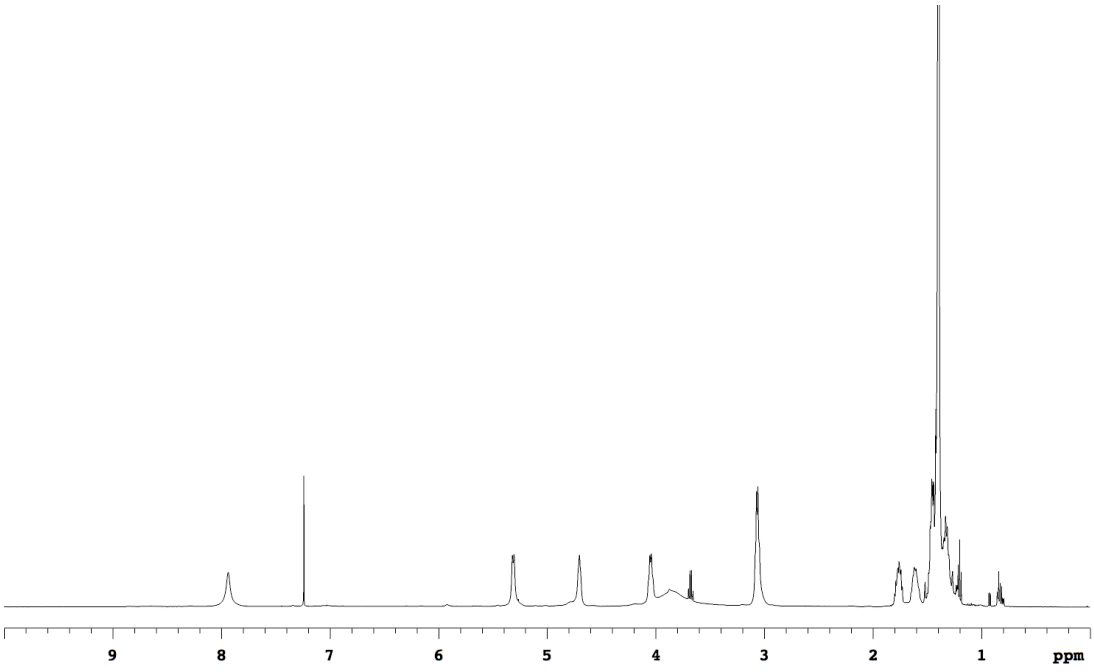


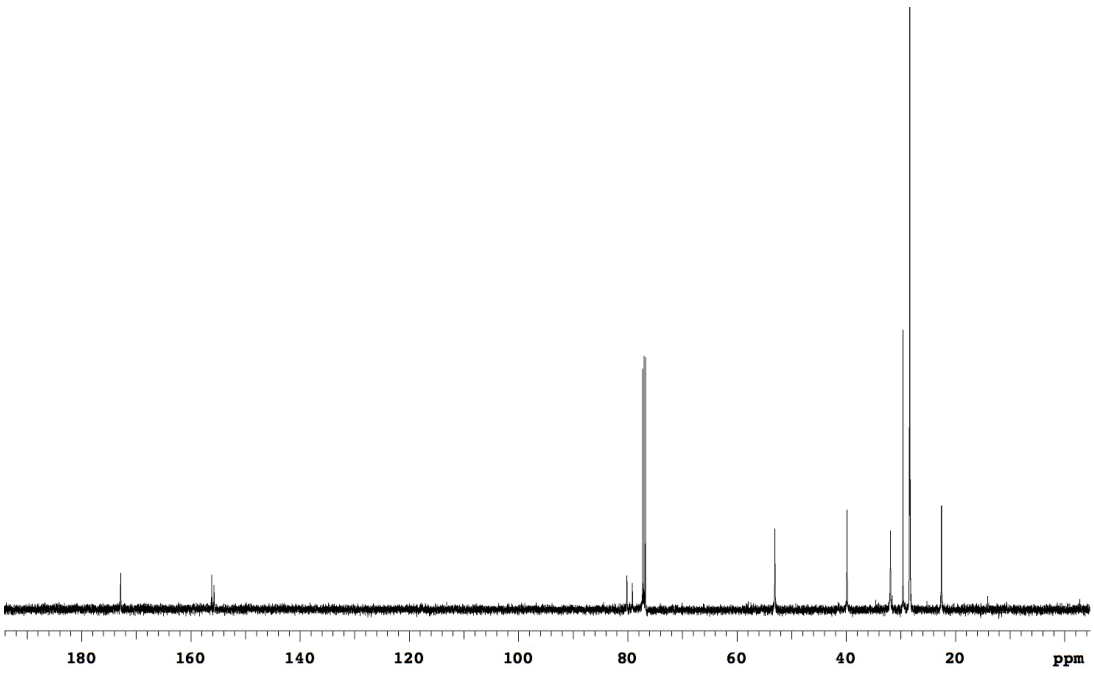
10c

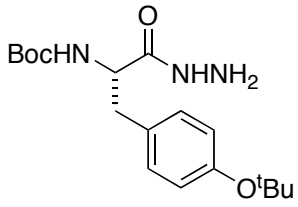
Compound **10c** was prepared from Boc-Lys(Boc)-OMe (5.41 g, 15.0 mmol) as a white solid.

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.94 (br, 1H), 5.31 (d, 1H, J = 7.5 Hz), 4.70 (br, 1H), 4.13-3.61 (m, 3H), 3.13-2.96 (m, 2H), 1.83-1.71 (m, 1H), 1.67-1.55 (m, 1H), 1.51-1.24 (m, 22H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 172.8, 156.2, 155.7, 80.1, 79.1, 53.0, 39.8, 31.9, 29.6, 28.4, 28.3, 22.5


9
8
7
6
5
4
3
2
1
ppm


180
160
140
120
100
80
60
40
20
ppm

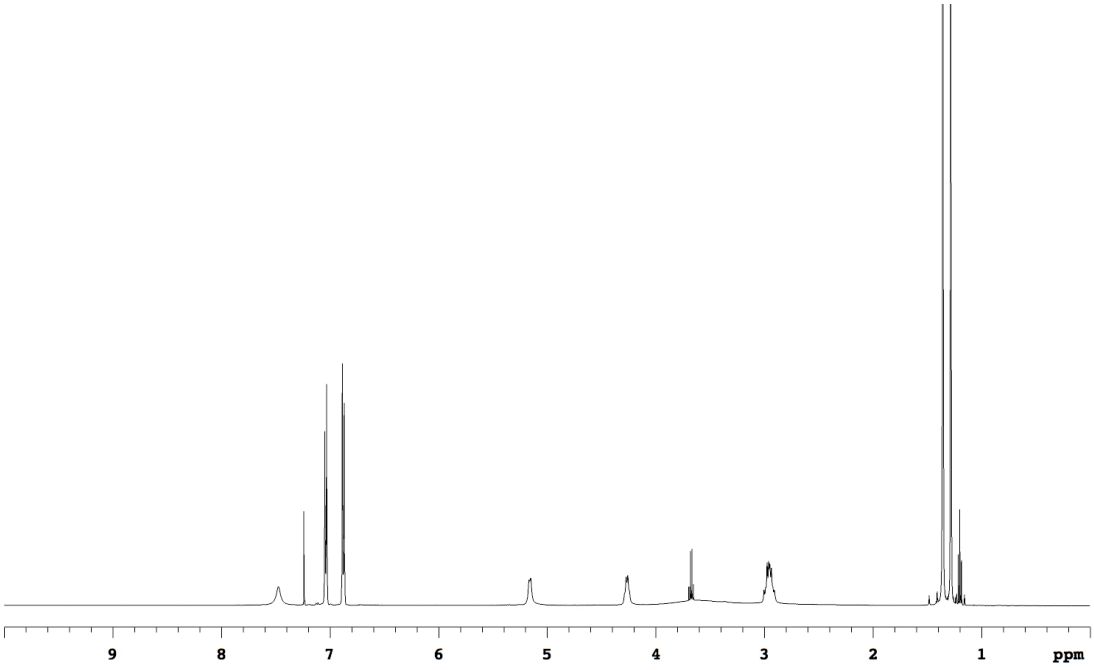


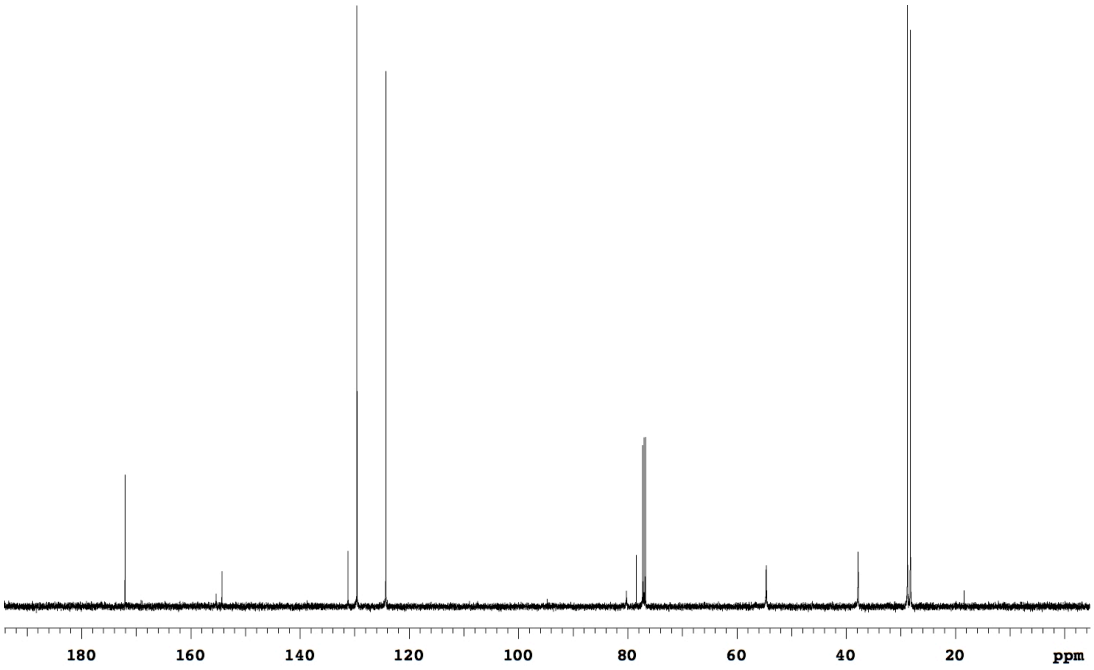
10d

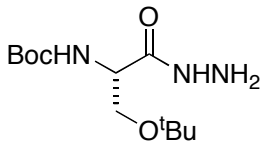
Compound **10d** was prepared from Boc-Tyr(^{t}Bu)-OMe (5.27 g, 15 .0 mmol) as a white solid.

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.48 (br, 1H), 7.04 (d, 2H, J = 6.9 Hz), 6.88 (d, 2H, J = 8.5 Hz), 5.16 (d, 1H, J = 7.5 Hz), 4.34-4.18 (m, 1H), 3.60 (br, 2H), 3.05-2.84 (m, 2H), 1.36 (s, 9H), 1.29 (s, 9H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 172.0, 155.4, 154.3, 131.2, 129.6, 124.3, 80.2, 78.4, 54.6, 37.8, 28.8, 28.2







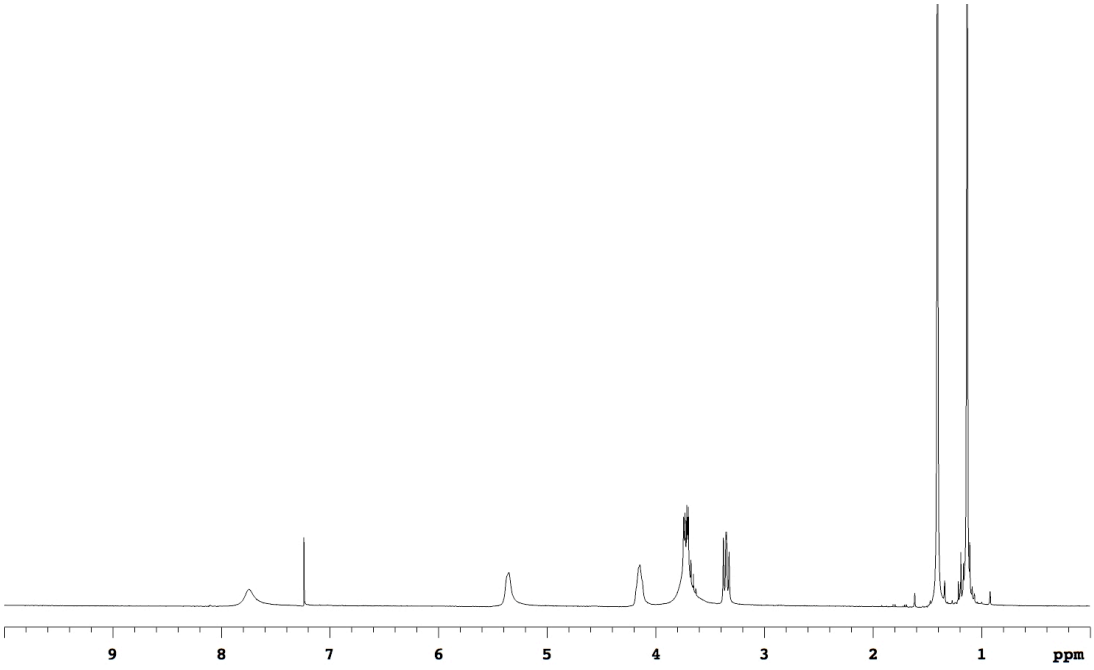
10e

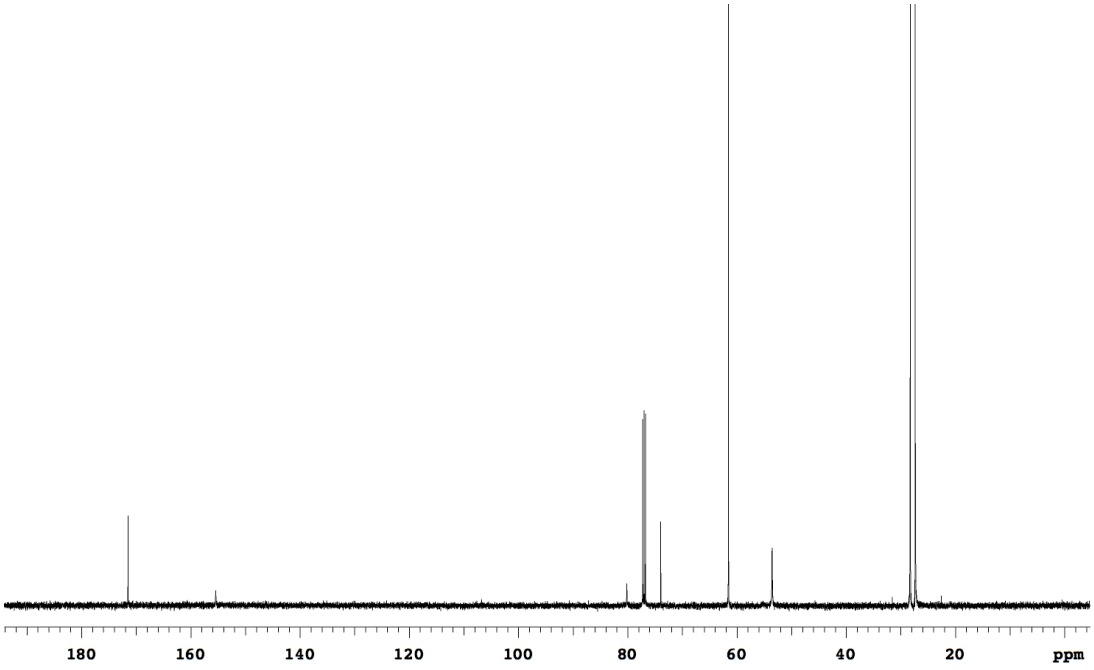
Compound **10e** was prepared from Boc-Ser(^{t}Bu)-OMe (4.13 g, 15.0 mmol) as a white solid.

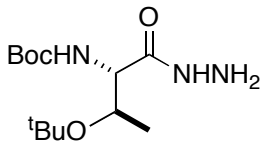
^{1}H NMR (300 MHz, $CDCl_3$) δ 7.74 (br, 1H), 5.35 (br, 1H), 4.15 (br, 1H), 3.88-3.49 (m, 3H), 3.35 (dd, 1H, J = 7.0, 8.7 Hz), 1.41 (s, 9H), 1.13 (s, 9H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 171.5, 155.4, 80.2, 73.9, 61.5, 53.6, 28.3, 27.3

MS (ESI, m/z) 276.2 $(M+H)^+$


9
8
7
6
5
4
3
2
1
ppm

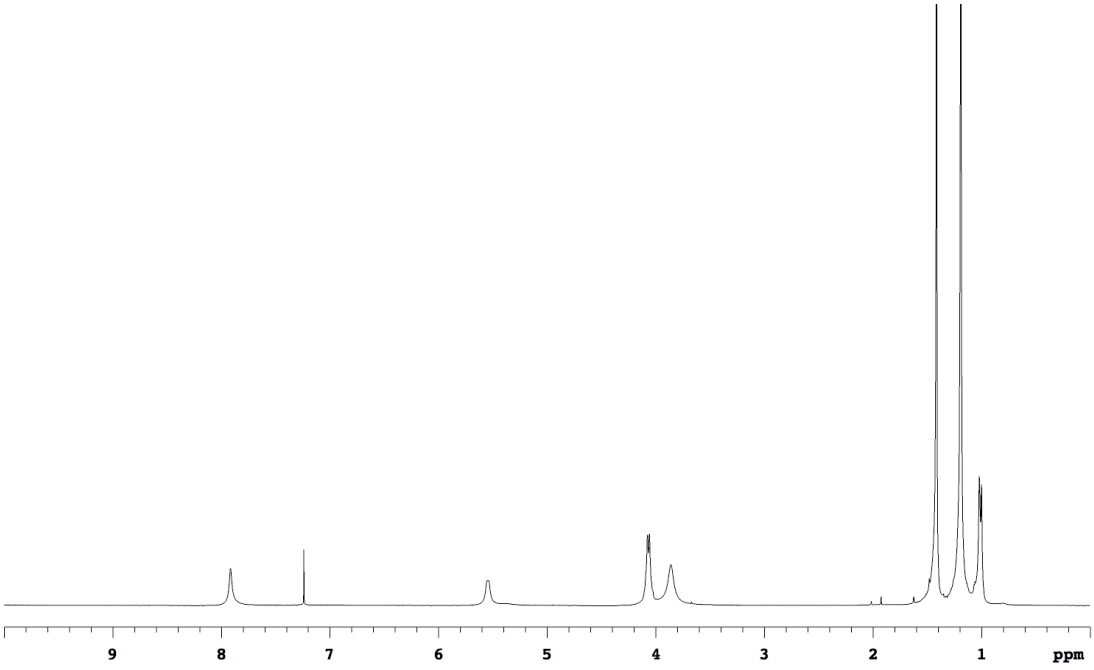

180
160
140
120
100
80
60
40
20
ppm

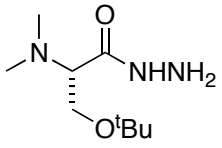


10f

Compound **10f** was prepared from Boc-Thr(^{t}Bu)-OMe (2.89 g, 10.0 mmol) as a white solid.

^{1}H NMR (300 MHz, $CDCl_3$) δ 7.91 (br, 1H), 5.54 (br, 1H), 4.20-3.99 (m, 2H), 3.86 (br, 2H), 1.41 (s, 9H), 1.19 (s, 9H), 1.01 (d, 3H, J = 6.0 Hz)





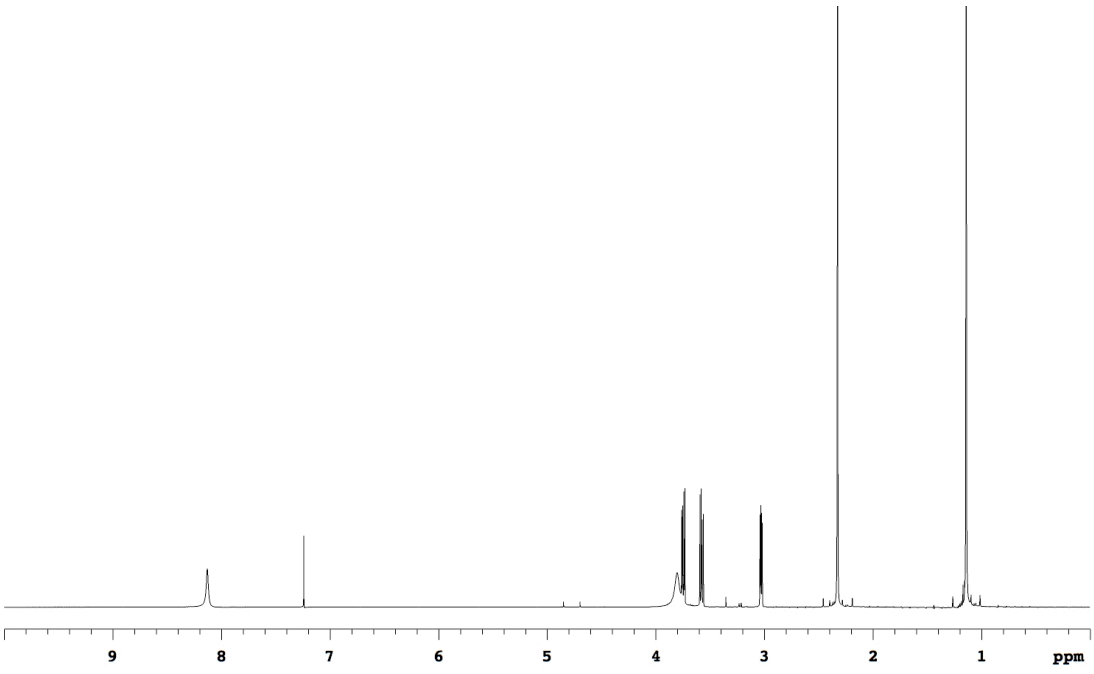
10g

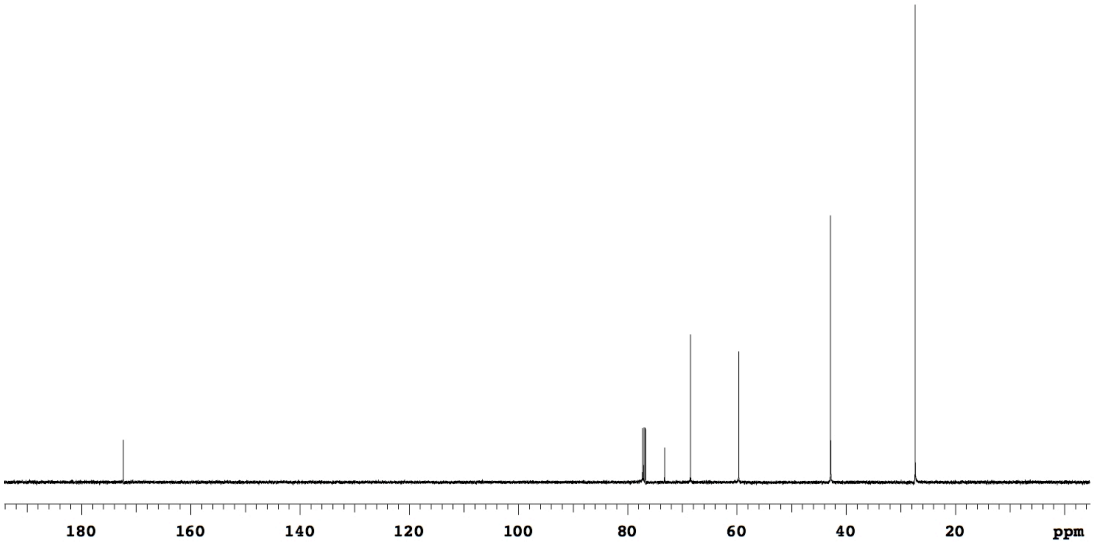
Compound **10g** was prepared from 13 (1.62 g, 8.0 mmol) *N*, *N*-dimethyl-Ser(^{t}Bu)-OMe as a white solid.

^{1}H NMR (500 MHz, $CDCl_3$) δ 8.13 (br, 1H), 3.80 (br, 2H), 3.75 (dd, 1H, J = 4.3, 9.8 Hz), 3.58 (dd, 1H, J = 6.1, 9.8 Hz), 3.03 (dd, 1H, J = 4.3, 6.1 Hz), 2.33 (s, 6H), 1.14 (s, 9H)

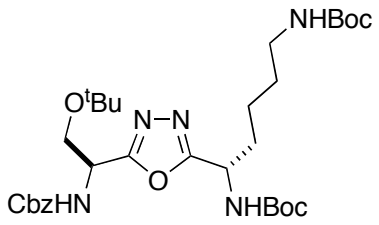
^{13}C NMR (125 MHz, $CDCl_3$) δ 172.3, 73.2, 68.5, 59.6, 42.8, 27.7

MS (APCI, m/z) 204.0 $(M+H)^+$





General Procedure for Preparation of 11a-q. The Cbz protected amino acid (1.0 equiv) was dissolved in dry CH_3CN (0.50 M). HOBt (1.1 equiv) was added in one portion followed by EDC (1.1 equiv). The mixture was stirred at room temperate for ~ 1h (all of the acid was converted to the activated ester/amide mixture). The resulting mixture was then slowly added to a solution of Boc protected amino acid hydrazide **10a-f** (1.0 equiv) in dry CH_3CN (0.50 M). The reaction mixture was vigorously stirred at room temperature for 8 h. The solvent was removed under reduced pressure and the residue was washed with H_2O (30 ml), 1N HCl (30 ml), saturated $NaHCO_3$ (30 ml) and brine (30 ml). The solid crude product was dried *in vacuo* and used in the next step directly. To a stirred solution of this diacyl hydrazide intermediate in dry THF (0.10 M), PPh_3 (2.0 equiv), I_2 (2.0 equiv) and Et_3N (4.0 equiv) was added at 0 °C. The cooling bath was removed after 30 min and stirring was continued at room temperature for 4-6 h. Water (30 ml) was added and stirring was continued for 1 h. The organic phase was separated and dried with Na_2SO_4 and concentrated *in vacuo*. The residue was purified by flash chromatography with EtOAc/hexanes mixtures to provide pure compounds **11a-q**.



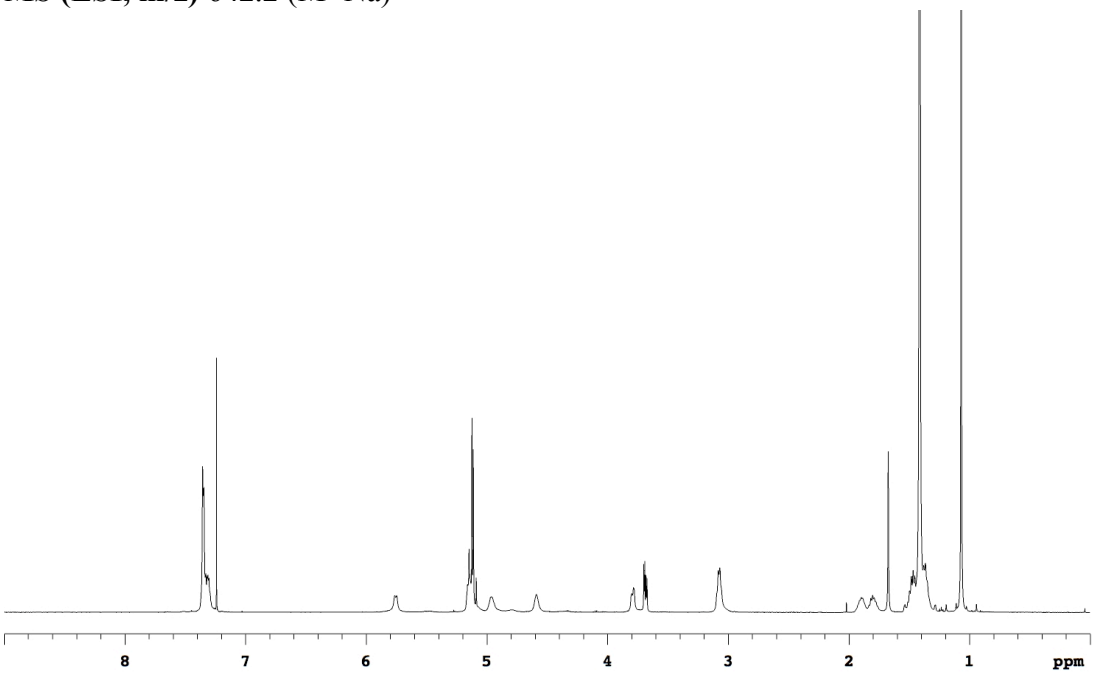
11a

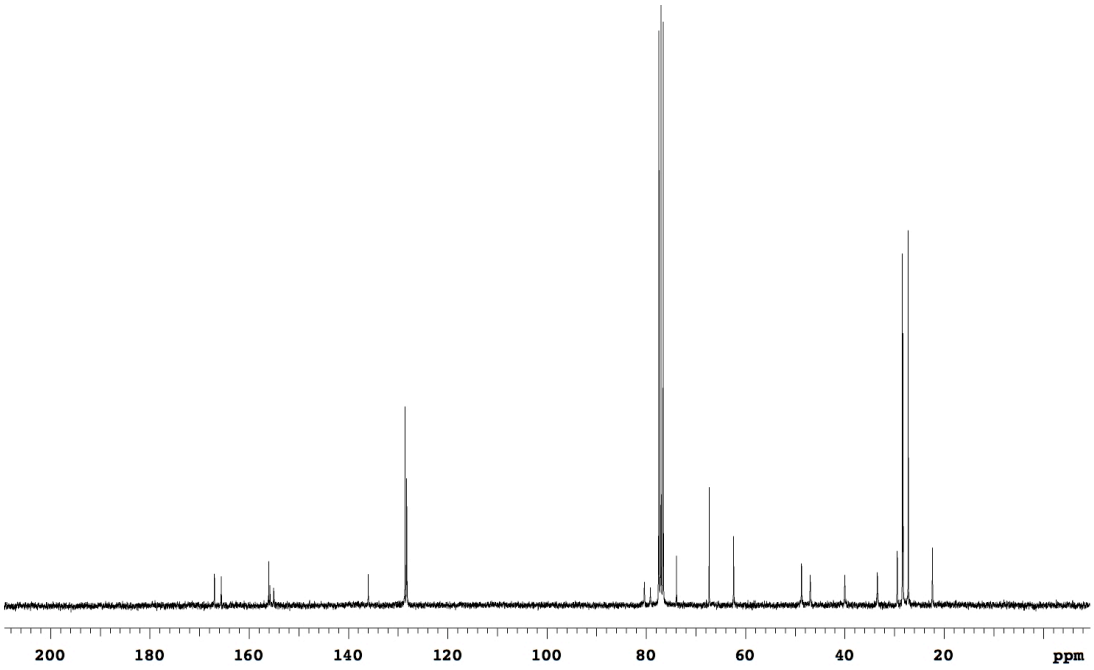
Compound **11a** was prepared from **10c** (4.75 g, 13.2 mmol) and Cbz-Ser(^{t}Bu)-OH (3.89 g, 13.2 mmol). Flash chromatography (20 % to 30 % EtOAc/Hexanes) afforded 5.27 g (64 %) **11a** as a white solid.

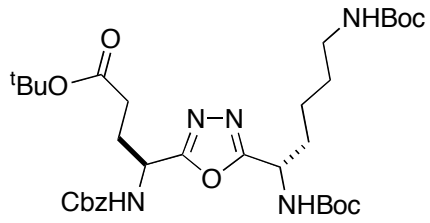
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.40-7.27 (m, 5H), 5.76 (d, 1H, J = 7.8 Hz), 5.19-5.07 (m, 4H), 4.96 (br, 1H), 4.59 (br, 1H), 3.82-3.76 (m, 1H), 3.69 (dd, 1H, J = 3.9, 9.2 Hz), 3.12-3.02 (m, 2H), 1.96-1.74 (m, 2H), 1.56-1.30 (m, 22H), 1.07 (s, 9H)

^{13}C NMR (75 MHz, $CDCl_3$) δ 167.0, 165.6, 156.0, 155.8, 155.0, 136.0, 128.6, 128.3, 128.2, 80.4, 79.2, 73.9, 67.3, 62.4, 48.7, 46.9, 40.0, 33.4, 29.4, 28.4, 28.3, 27.2, 22.3

MS (ESI, m/z) 642.2 $(M+Na)^+$







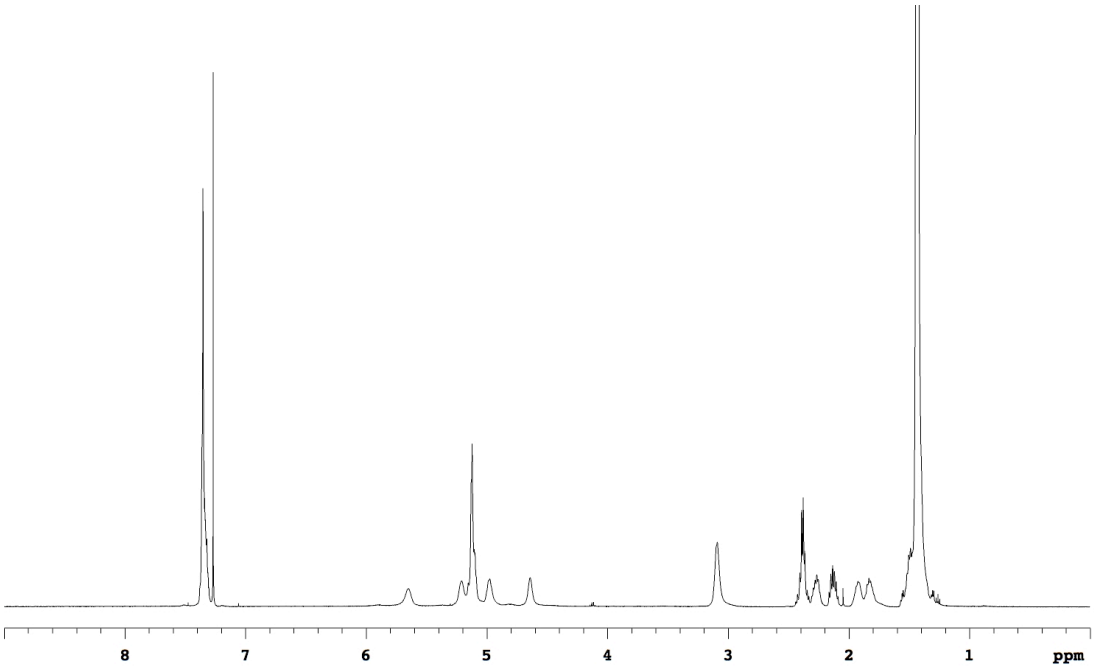
11b

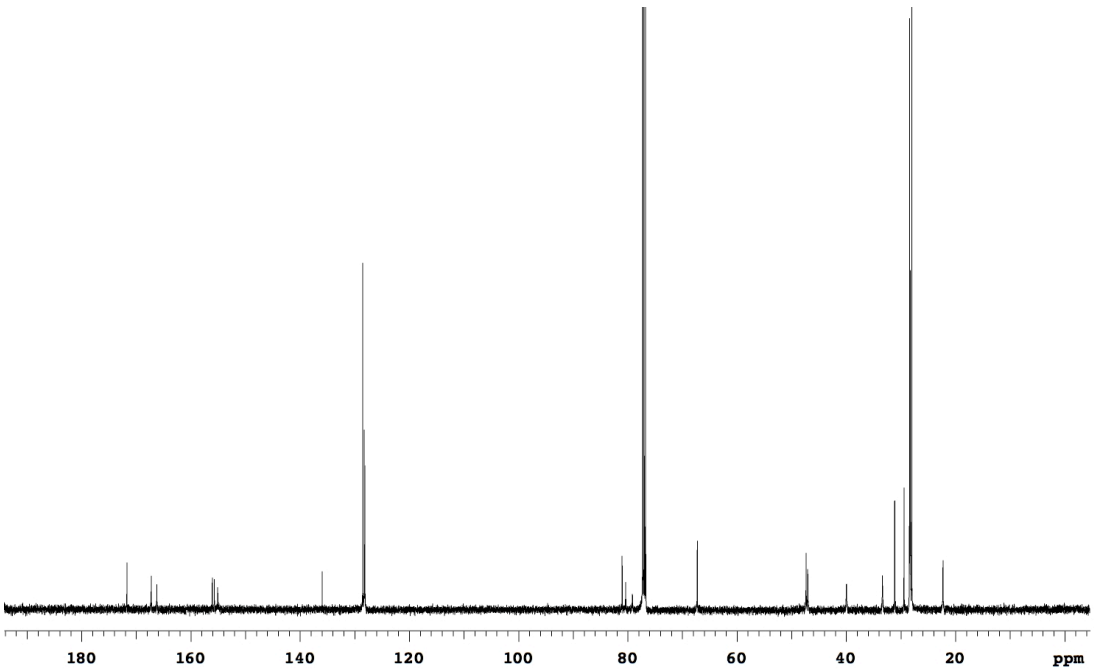
Compound **11b** was prepared from **10c** (3.48 g, 14.2 mmol) and Cbz-Glu(O^{t}Bu)-OH (5.40 g, 14.2 mmol). Flash chromatography (20 % to 30 % EtOAc/Hexanes) afforded 4.52 g (48 %) **11b** as a white solid.

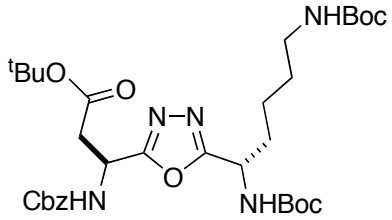
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.40-7.29 (m, 5H), 5.65 (br, 1H), 5.27-5.05 (m, 4H), 4.97 (br, 1H), 4.64 (br, 1H), 3.09 (br, 2H), 2.45-2.21 (m, 3H), 1.18-1.08 (m, 1H), 1.58-1.33 (m, 31H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 171.7, 167.3, 166.2, 156.1, 155.7, 155.1, 135.9, 128.5, 128.3, 128.1, 81.0, 80.4, 79.1, 67.3, 47.4, 47.0, 39.9, 33.4, 31.1, 29.4, 28.4, 28.3, 28.0, 22.3

MS (ESI, m/z) 684.3 $(M+Na)^+$


8
7
6
5
4
3
2
1
ppm


180
160
140
120
100
80
60
40
20
ppm



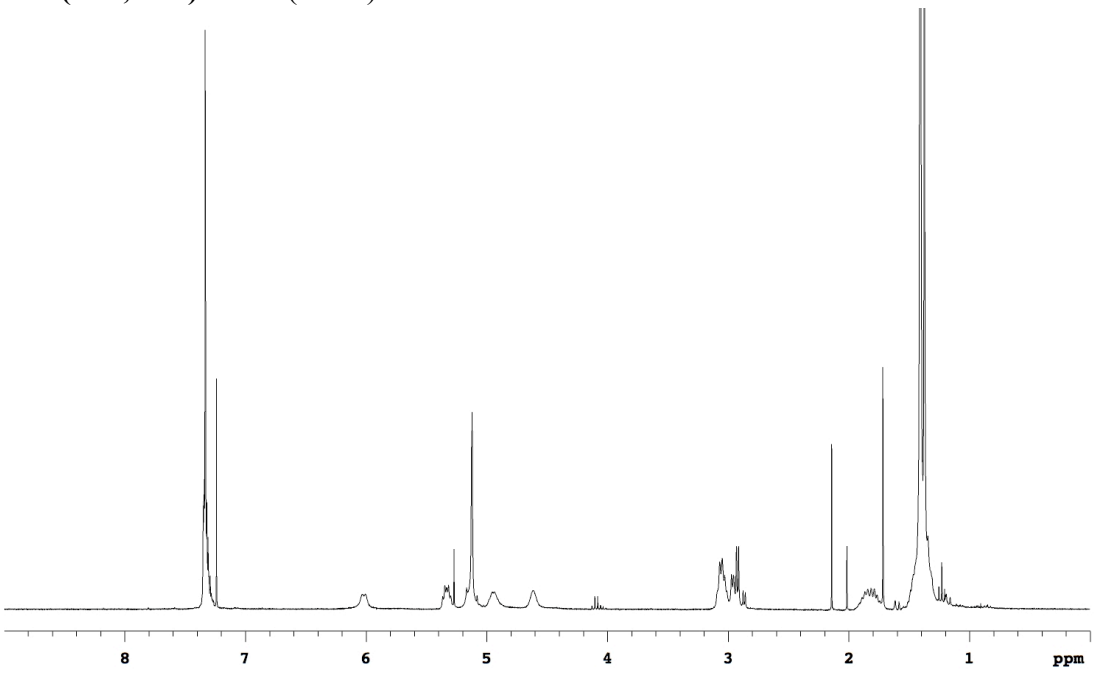
11c

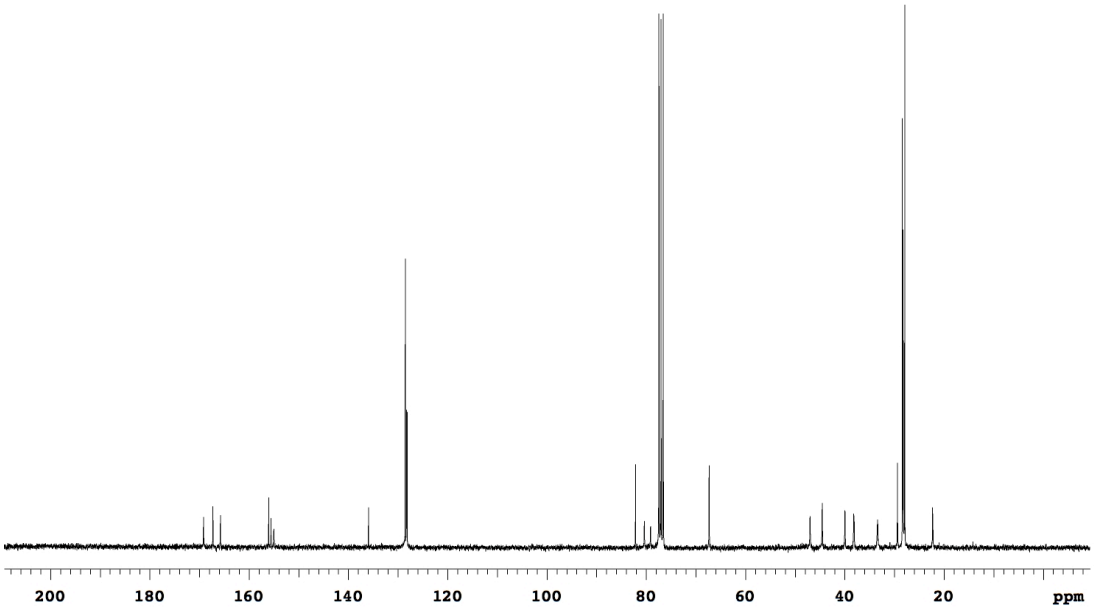
Compound **11c** was prepared from **10c** (4.66 g, 12.9 mmol) and Cbz-Asp(O^{t}Bu)-OH•DCHA (6.53 g, 12.9 mmol). Flash chromatography (20 % to 30 % EtOAc/Hexanes) afforded 3.32 g (40 %) **11c** as a white solid.

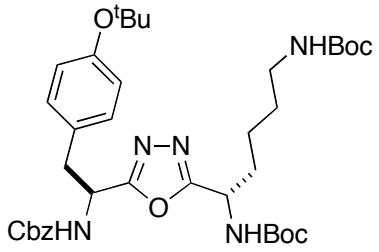
^{1}H NMR (300 MHz, $CDCl_3$) δ 7.39-7.26 (m, 5H), 6.02 (d, 1H, J = 7.8 Hz), 5.39-5.28 (m, 1H), 5.21-5.06 (m, 3H), 4.95 (br, 1H), 4.61 (br, 1H), 3.12-2.84 (m, 4H), 1.95-1.73 (m, 2H), 1.53-1.27 (m, 31H)

^{13}C NMR (75 MHz, $CDCl_3$) δ 169.2, 167.3, 165.8, 156.0, 155.6, 155.0, 135.9, 128.5, 128.3, 128.2, 82.1, 80.4, 79.1, 67.3, 47.0, 44.5, 40.0, 38.2, 33.4, 29.4, 28.4, 28.2, 27.9, 22.3

MS (ESI, m/z) 648.4 $(M+H)^+$







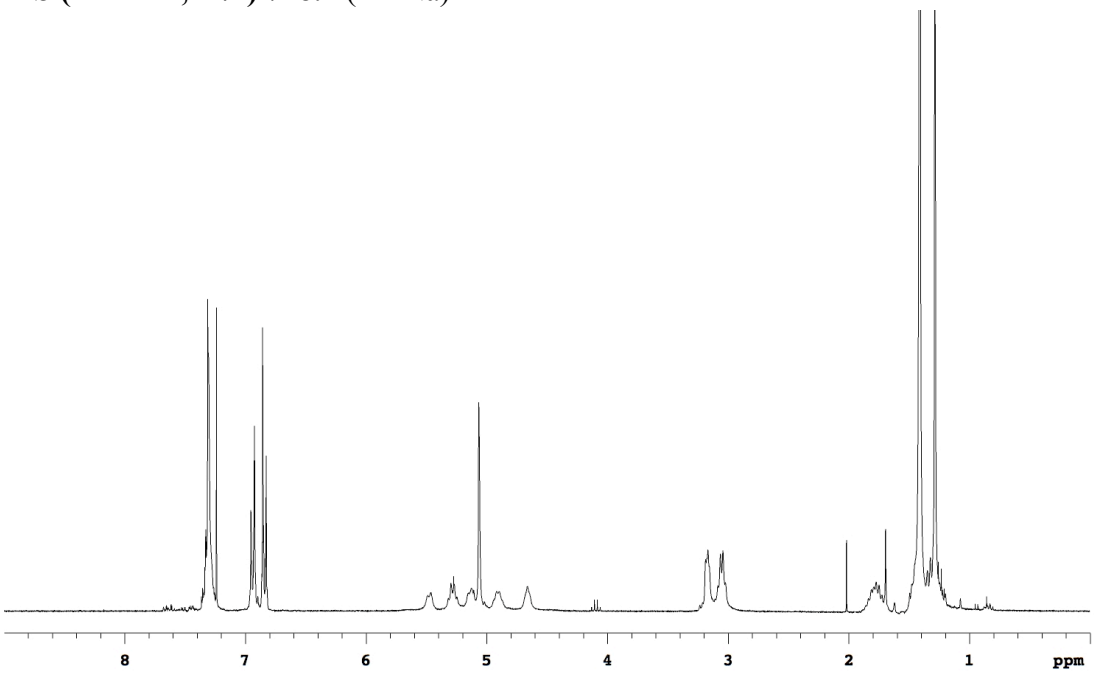
11d

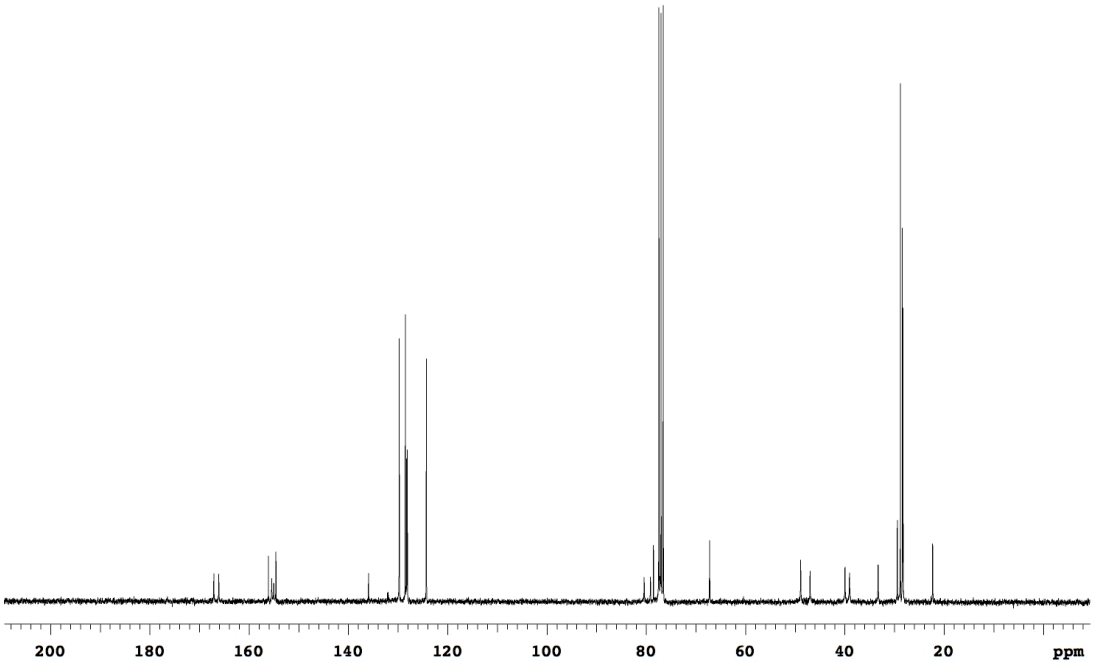
Compound **11d** was prepared from **10c** (4.70 g, 13.0 mmol) and Cbz-Tyr(^{t}Bu)-OH (4.84 g, 13.0 mmol). Flash chromatography (20 % to 30 % EtOAc/Hexanes) afforded 3.70 g (41 %) **11d** as a white solid.

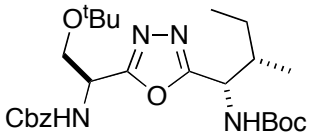
^{1}H NMR (300 MHz, $CDCl_3$) δ 7.39-7.26 (m, 5H), 6.94 (d, 2H, J = 8.5 Hz), 6.84 (d, 2H, J = 8.5 Hz), 5.48 (d, 1H, J = 8.0 Hz), 5.36-5.21 (m, 1H), 5.20-5.00 (m, 3H), 4.97-4.84 (m, 1H), 4.66 (br, 1H), 3.25-2.96 (m, 4H), 1.90-1.66 (m, 2H), 1.52-1.19 (m, 31H)

^{13}C NMR (75 MHz, $CDCl_3$) δ 167.1, 166.1, 156.1, 155.4, 155.0, 154.6, 135.9, 132.1, 129.7, 128.5, 128.3, 128.1, 124.3, 80.4, 79.1, 78.5, 67.2, 48.9, 47.0, 40.0, 39.0, 33.3, 29.4, 28.8, 28.4, 28.3, 22.3

MS (MALDI, m/z) 718.1 $(M+Na)^+$







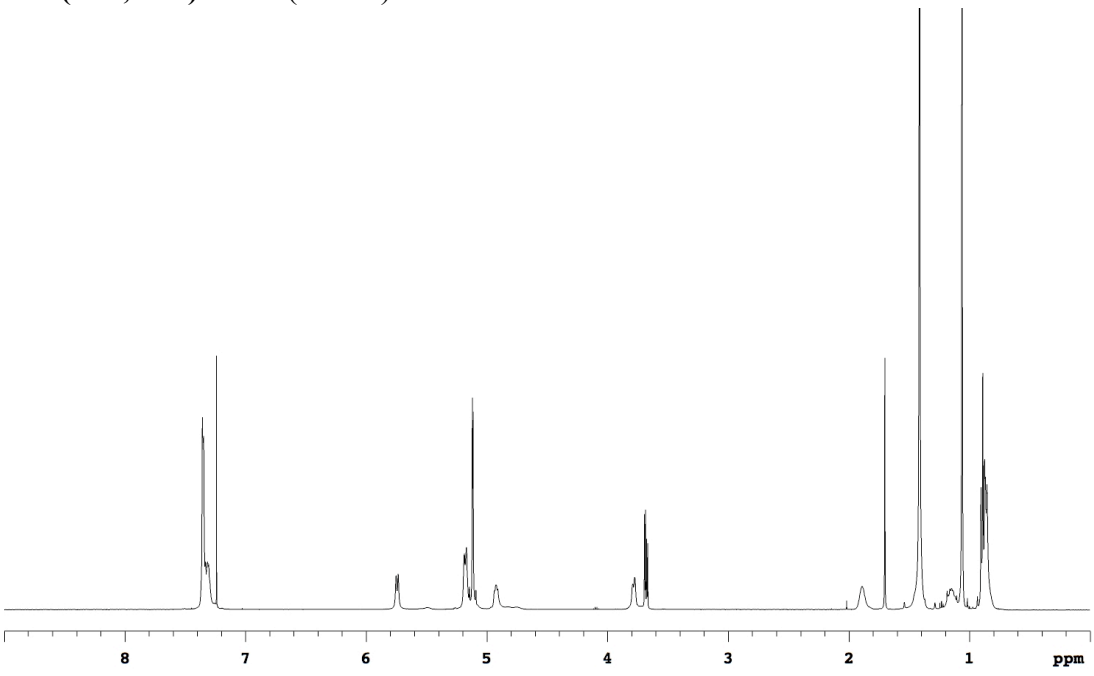
11e

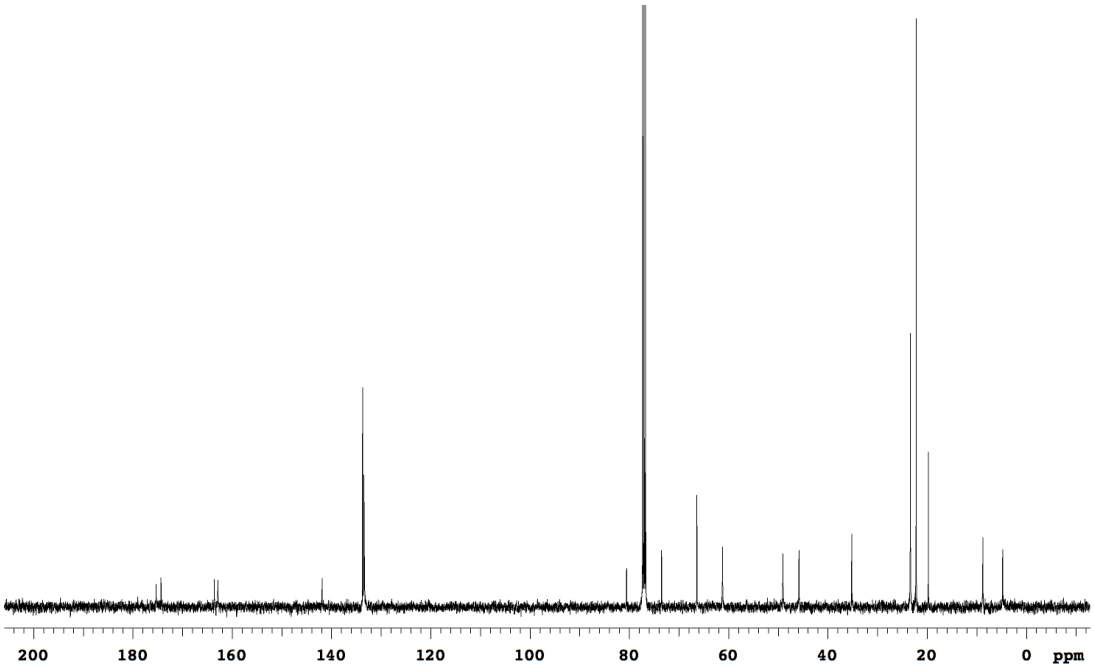
Compound **11e** was prepared from **10b** (3.51 g, 14.3 mmol) and Cbz-Ser(^{t}Bu)-OH (4.22 g, 14.3 mmol). Flash chromatography (20 % to 30 % EtOAc/Hexanes) afforded 5.42 g (75 %) **11e** as a white solid.

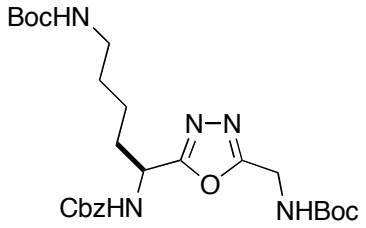
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.40-7.27 (m, 5H), 5.74 (d, 1H, J = 8.8 Hz), 5.23-5.06 (m, 4H), 4.96-4.88 (m, 1H), 3.83-3.75 (m, 1H), 3.68 (dd, 1H, J = 3.9, 9.1 Hz), 1.89 (br, 1H), 1.48-1.35 (m, 10H), 1.21-1.10 (m, 1H), 1.06 (s, 9H), 0.93-0.80 (m, 6H)

^{13}C NMR (75 MHz, $CDCl_3$) δ 175.3, 174.3, 163.6, 162.9, 141.9, 133.7, 133.4, 133.4, 80.5, 73.5, 66.4, 61.2, 49.1, 45.8, 35.2, 23.4, 22.2, 19.7, 8.8, 4.8

MS (ESI, m/z) 527.3 $(M+Na)^+$







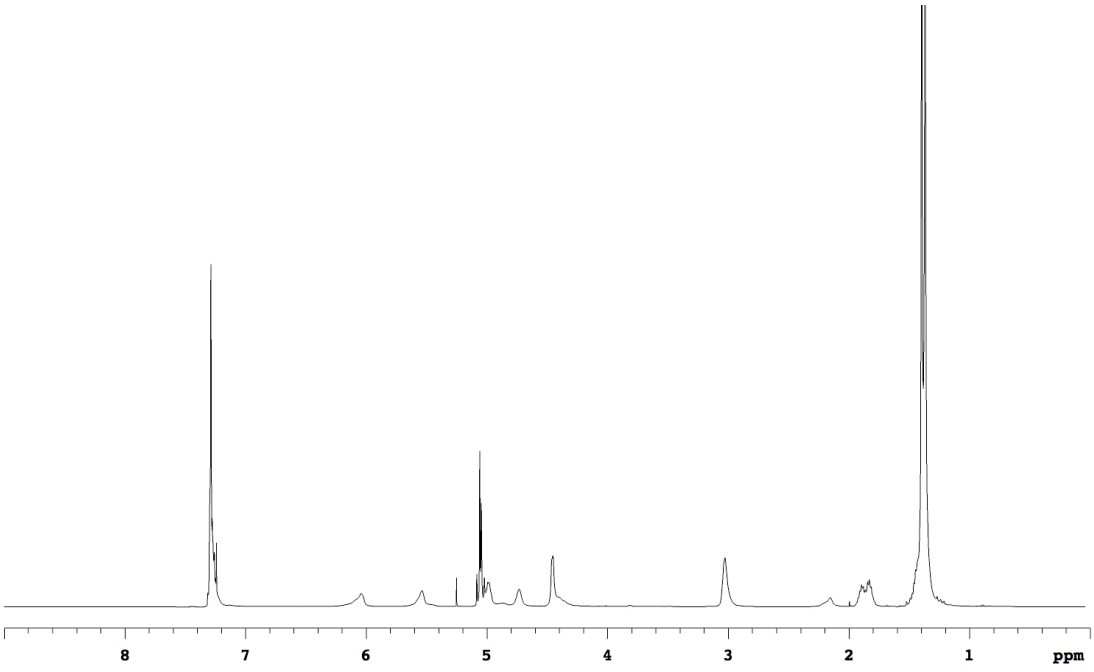
11f

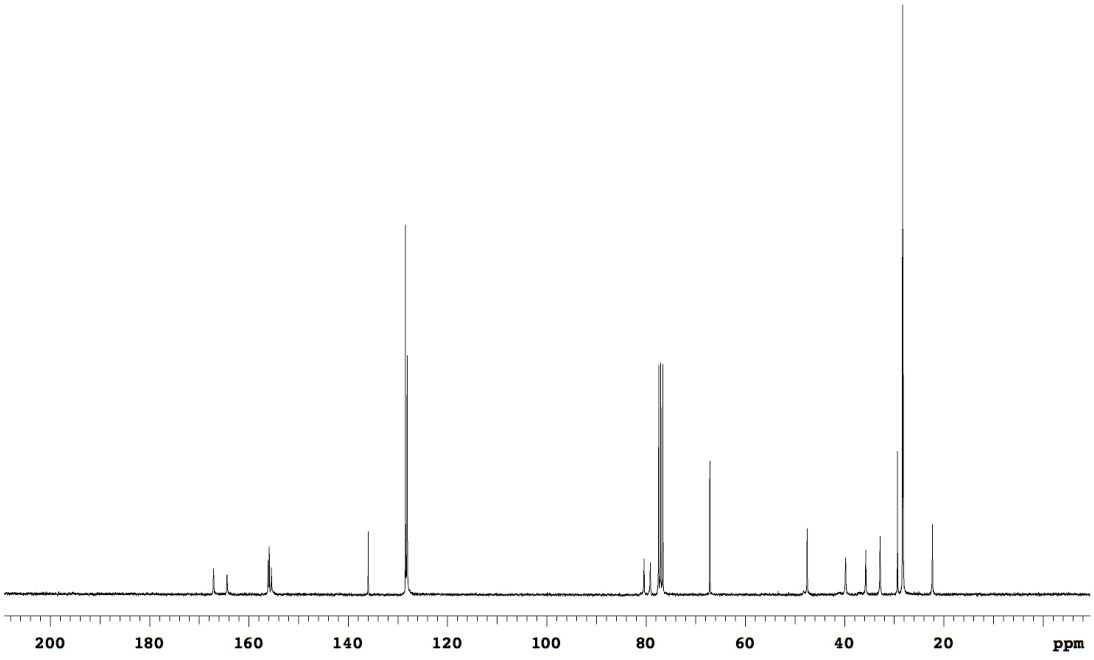
Compound **11f** was prepared from **10a** (3.37 g, 17.8 mmol) and Cbz-Lys(Boc)-OH (6.78 g, 17.8 mmol). Flash chromatography (25 % to 50 % EtOAc/Hexanes) afforded 5.57 g (59%) **11f** as a white solid.

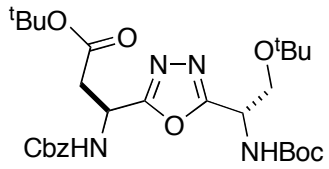
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.34-7.18 (m, 5H), 6.04 (br, 1H), 5.54 (br, 1H), 5.10-4.94 (m, 3H), 4.73 (br, 1H), 4.50-4.30 (br, 2H), 3.03 (br, 2H), 1.96-1.75 (m, 2H), 1.53-1.20 (m, 22H)

^{13}C NMR (75 MHz, $CDCl_3$) δ 167.1, 164.4, 156.1, 155.9, 155.4, 135.9, 128.4, 128.2, 128.1, 80.4, 79.1, 67.1, 47.5, 39.8, 35.7, 32.8, 29.3, 28.3, 28.2, 22.3

MS (ESI, m/z) 556.3 $(M+Na)^+$


8
7
6
5
4
3
2
1
ppm


200
180
160
140
120
100
80
60
40
20
ppm



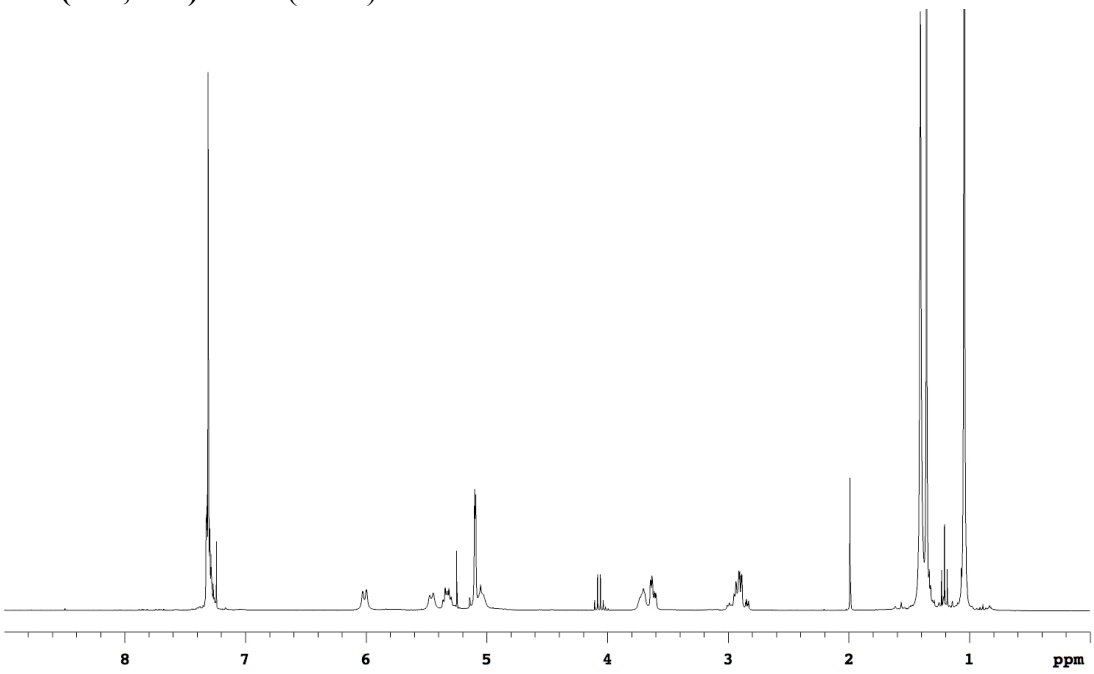
11g

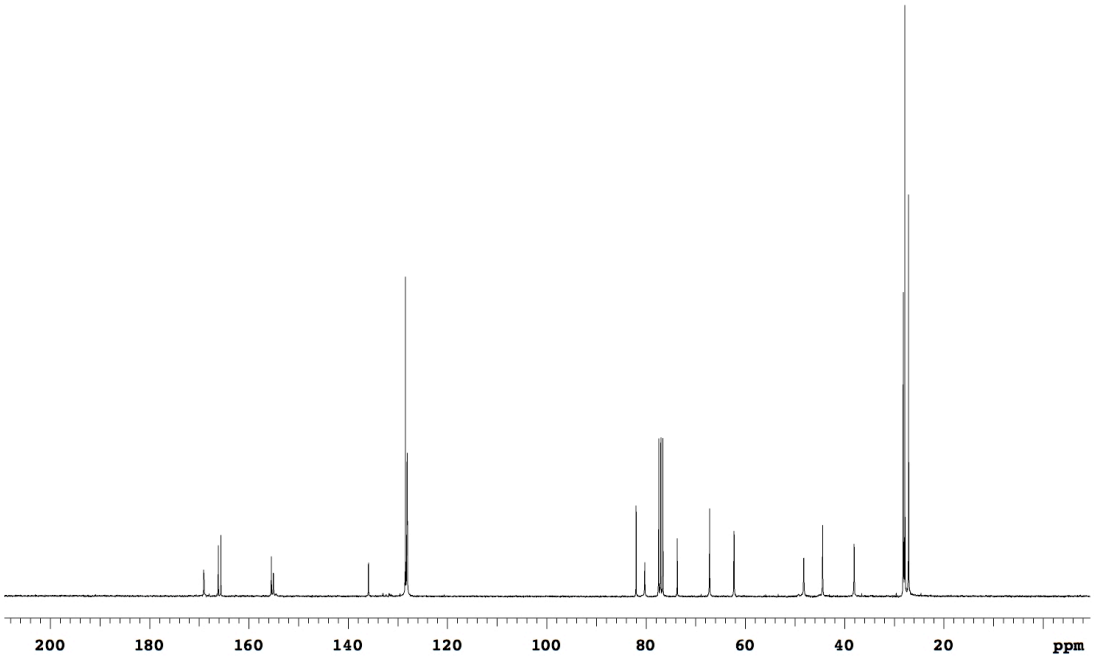
Compound **11g** was prepared from **10e** (4.13 g, 15.0 mmol) and Cbz-Asp(O^{t}Bu)-OH•DCHA (7.57 g, 15.0 mmol). Flash chromatography (20 % to 30 % EtOAc/Hexanes) afforded 3.80 g (45 %) **11g** as a white solid.

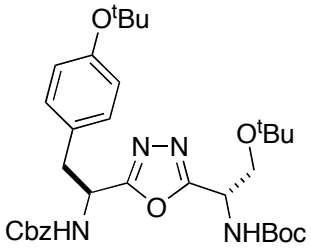
^{1}H NMR (300 MHz, $CDCl_3$) δ 7.36-7.25 (m, 5H), 6.01 (d, 1H, J = 9.2 Hz), 5.46 (d, 1H J = 8.7 Hz), 5.38-5.28 (m, 1H), 5.16-4.98 (m, 3H), 3.78-3.67 (m, 1H), 3.62 (dd, 1H, J = 4.0, 9.2 Hz), 3.03-2.81 (m, 2H), 1.47-1.29 (m, 18H), 1.04 (s, 9H)

^{13}C NMR (75 MHz, $CDCl_3$) δ 169.1, 166.2, 165.6, 155.5, 155.0, 135.9, 128.2, 128.1, 128.0, 82.0, 80.2, 73.7, 67.2, 62.2, 48.2, 44.4, 38.0, 28.2, 27.8, 27.1

MS (ESI, m/z) 563.3 $(M+H)^+$





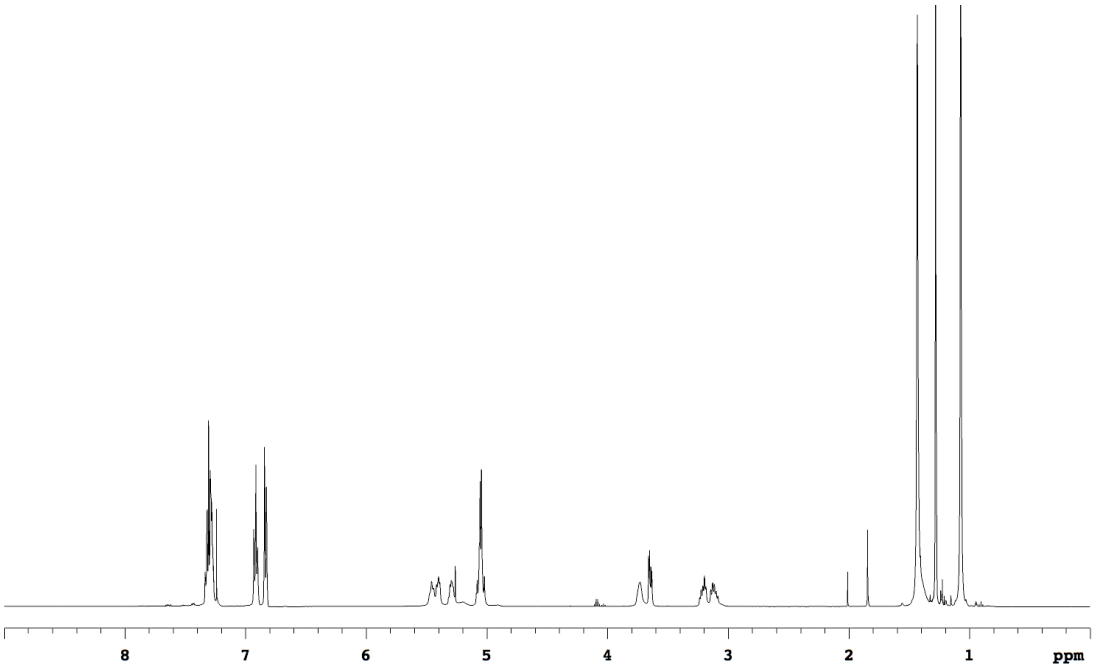


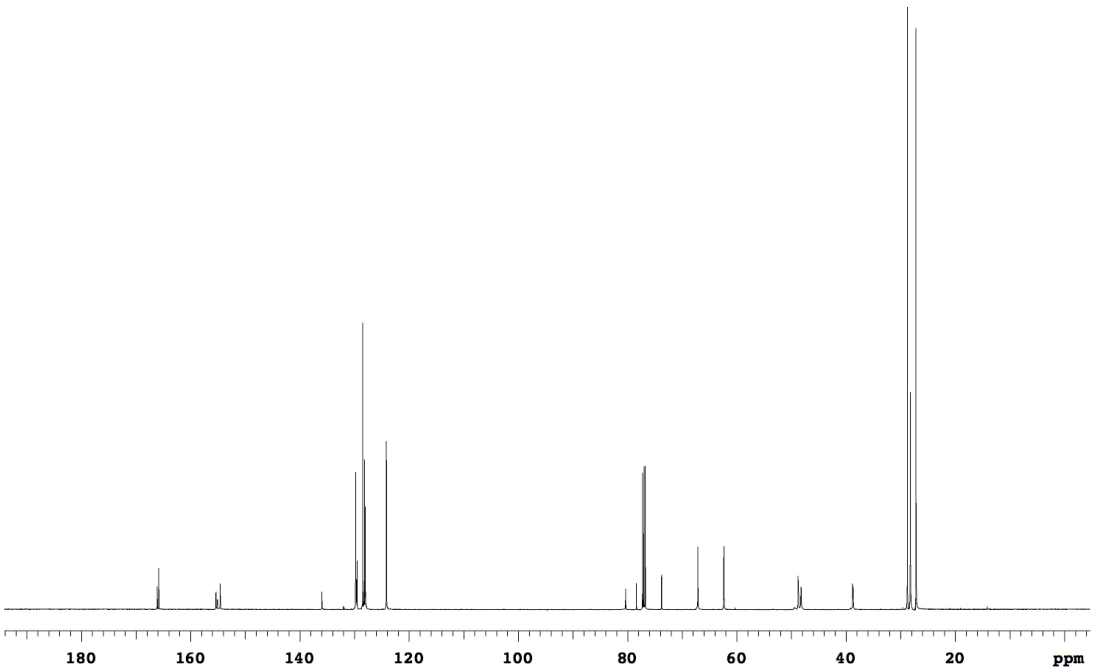
11h

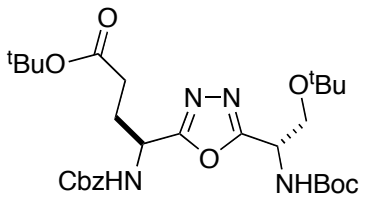
Compound **11h** was prepared from **10e** (2.34 g, 8.5 mmol) and Cbz-Tyr(^{t}Bu)-OH•DCHA (4.70 g, 8.5 mmol). Flash chromatography (20 % to 30 % EtOAc/Hexanes) afforded 4.20 g (81%) **11h** as a white solid.

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.35-7.26 (m, 5H), 6.95-6.88 (m, 2H), 6.86-6.81 (m, 2H), 5.50-5.37 (m, 2H), 5.33-5.25 (m, 1H), 5.10-5.00 (m, 3H), 3.73 (br, 1H), 3.64 (dd, 1H, J = 4.0, 9.2 Hz), 3.26-3.02 (m, 2H), 1.43 (s, 9H), 1.28 (s, 9H), 1.07 (s, 9H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 166.1, 165.8, 155.4, 155.1, 154.6, 136.0, 129.8, 129.5, 128.5, 128.2, 128.0, 124.1, 80.4, 78.4, 73.8, 67.1, 62.3, 48.8, 48.2, 38.8, 28.8, 28.2, 27.2


8
7
6
5
4
3
2
1
ppm


180
160
140
120
100
80
60
40
20
ppm



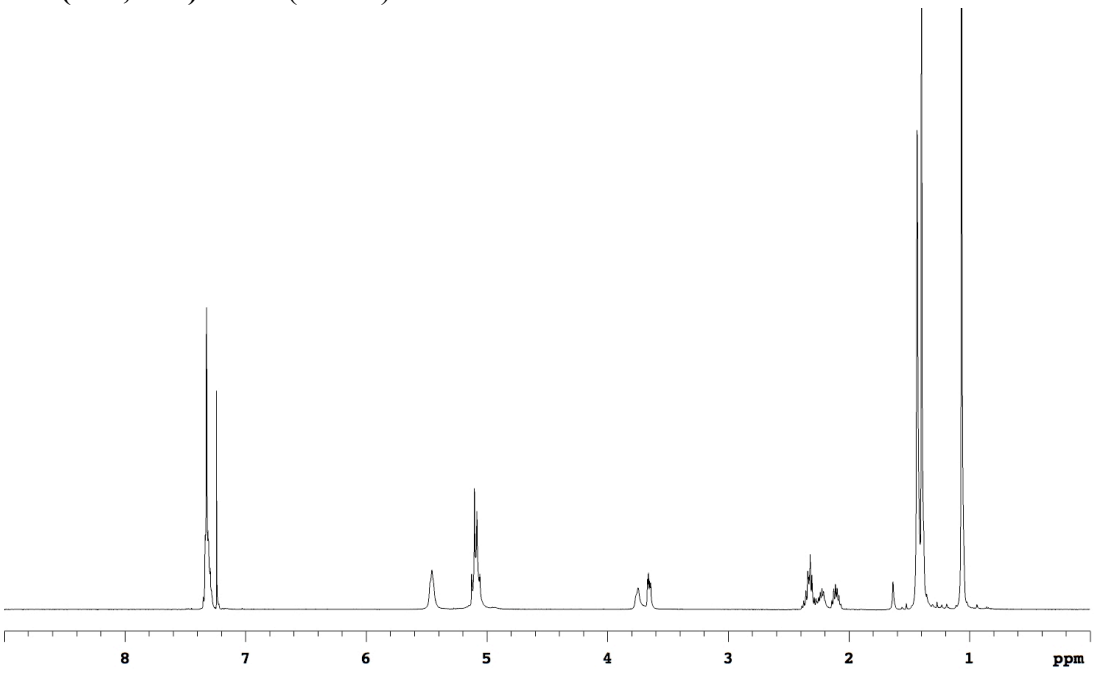
11i

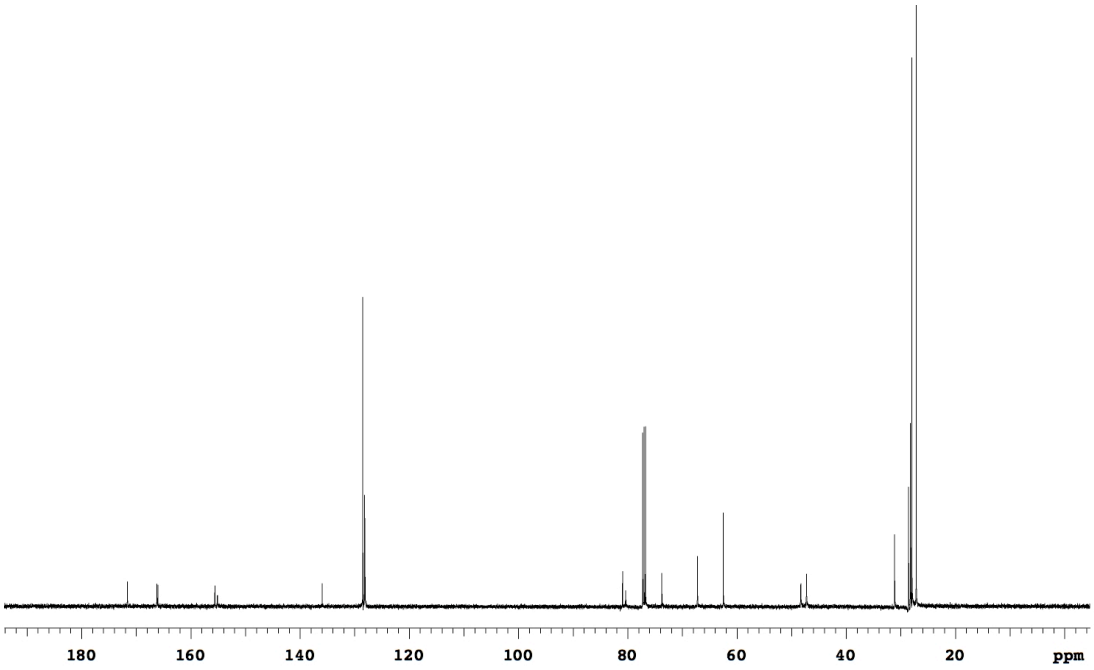
Compound **11i** was prepared from **10e** (3.99 g, 14.5 mmol) and Cbz-Glu(O^{t}Bu)-OH (4.89 g, 14.5 mmol). Flash chromatography (20 % to 30 % EtOAc/Hexanes) afforded 7.02 g (48 %) **11i** as a white solid.

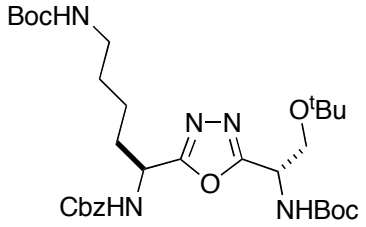
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.36-7.26 (m, 5H), 5.45 (br, 2H), 5.15-5.02 (m, 4H), 3.80-3.70 (br, 1H), 3.65 (dd, 1H, J = 3.8, 9.1 Hz), 2.41-2.17 (m, 3H), 2.16-2.05 (m, 1H), 1.49-1.32 (m, 18H), 1.07 (s, 9H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 171.6, 166.2, 166.0, 155.6, 155.1, 135.9, 128.5, 128.2, 128.1, 80.9, 80.4, 73.7, 67.2, 62.5, 48.3, 47.3, 31.1, 28.6, 28.2, 28.0, 27.2

MS (ESI, m/z) 599.3 $(M+Na)^+$







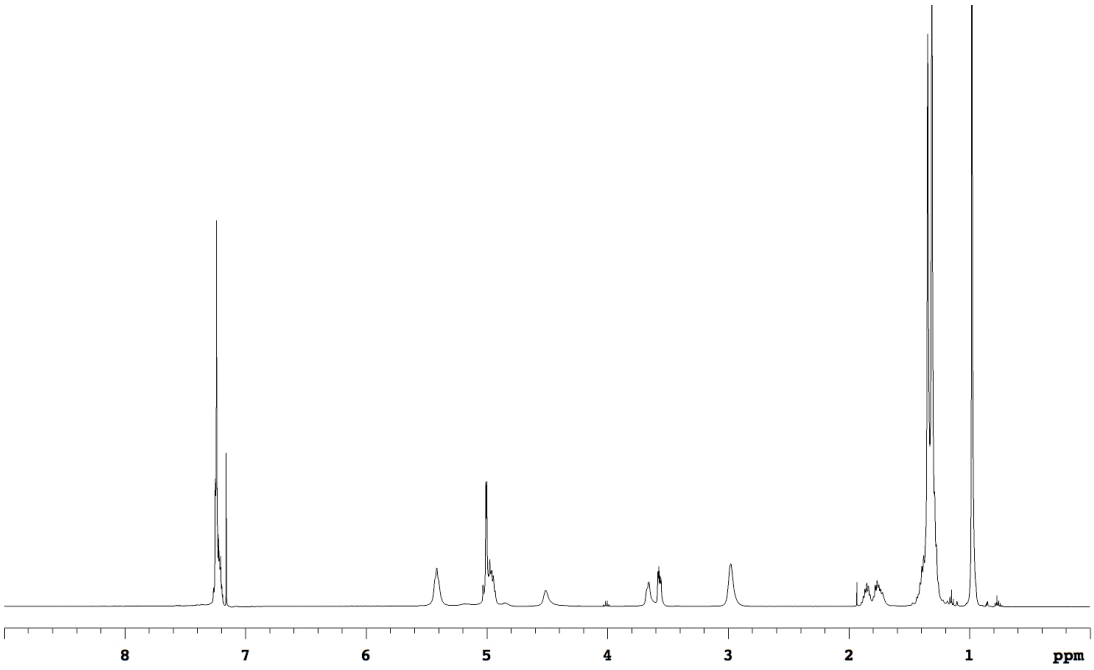
11j

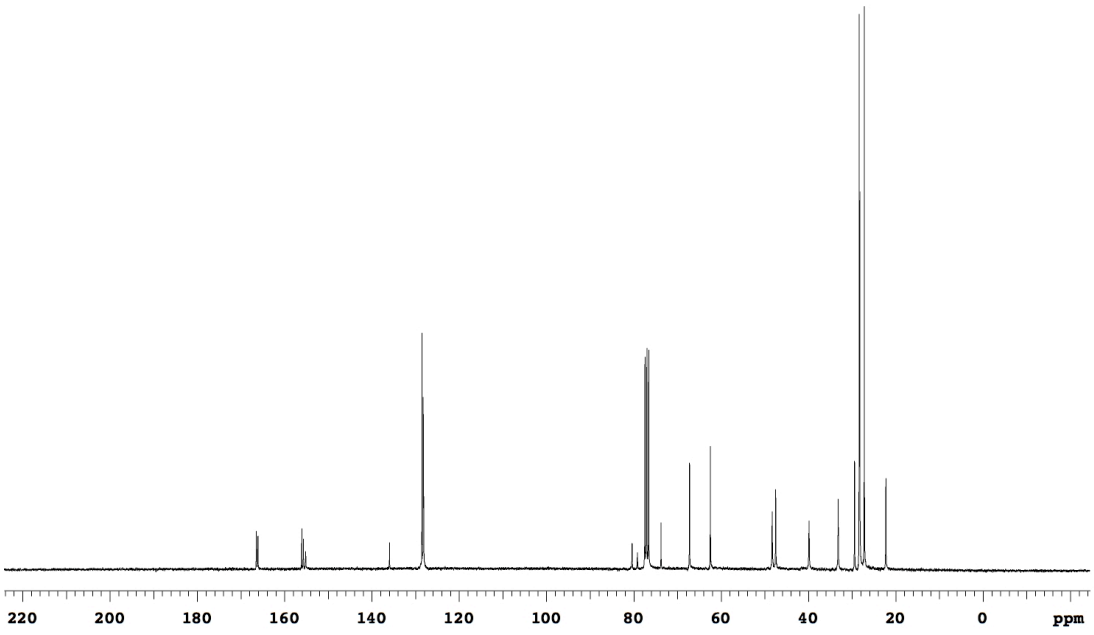
Compound **11j** was prepared from **10e** (3.71 g, 13.5 mmol) and Cbz-Lys(Boc)-OH (4.55 g, 13.5 mmol). Flash chromatography (20 % to 30 % EtOAc/Hexanes) afforded 5.0 g (60%) **11j** as a white solid.

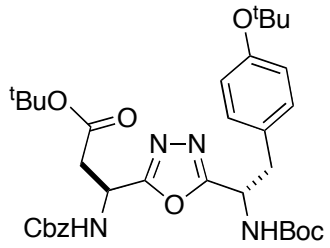
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.39-7.26 (m, 5H), 5.50 (br, 2H), 5.15-4.97 (m, 4H), 4.59 (br, 1H), 3.74 (br, 1H), 3.65 (dd, 1H, J = 3.9, 9.1 Hz), 3.06 (br, 2H), 1.99-1.75 (m, 2H), 1.54-1.28 (m, 22H), 1,06 (s, 9H)

^{13}C NMR (75 MHz, $CDCl_3$) δ 166.5, 166.1, 156.0, 155.7, 155.2, 135.9, 128.5, 128.2, 128.1, 80.4, 79.2, 73.8, 67.2, 62.5, 48.3, 47.5, 39.9, 33.2, 29.4, 28.4, 28.2, 27.2, 22.2

MS (ESI, m/z) 642.4 $(M+Na)^+$


8
7
6
5
4
3
2
1
ppm


220
200
180
160
140
120
100
80
60
40
20
0
ppm



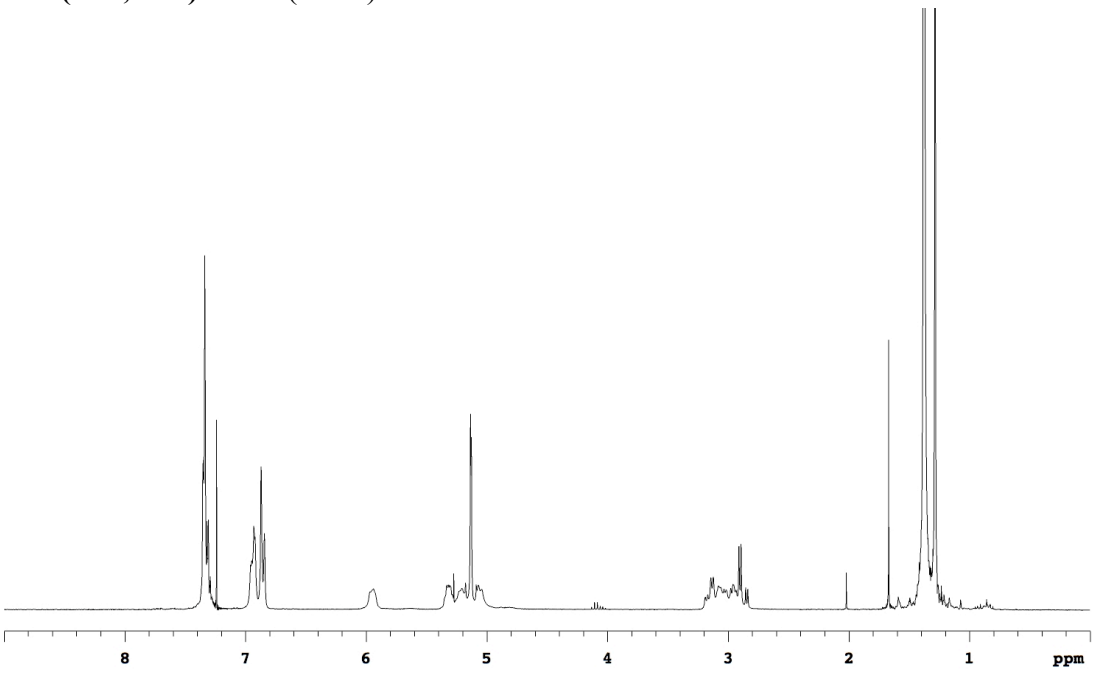
11k

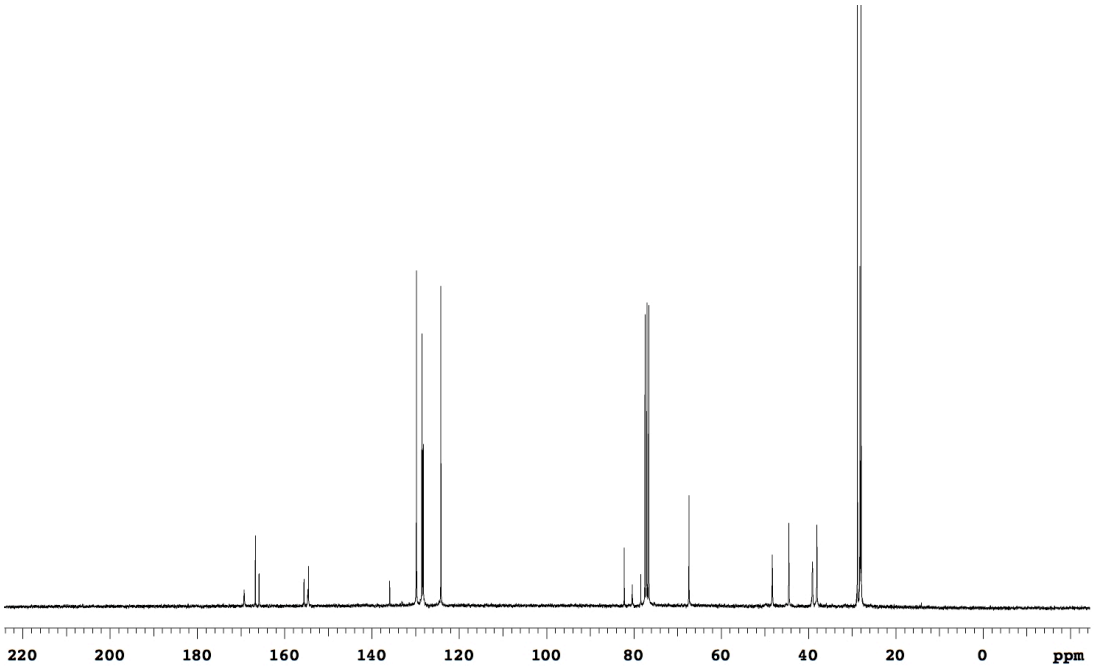
Compound **11k** was prepared from **10d** (4.43 g, 12.6 mmol) and Cbz-Asp(O^{t}Bu)-OH•DCHA (6.36 g, 12.6 mmol). Flash chromatography (20 % to 30 % EtOAc/Hexanes) afforded 3.40 g (42%) **11k** as a white solid.

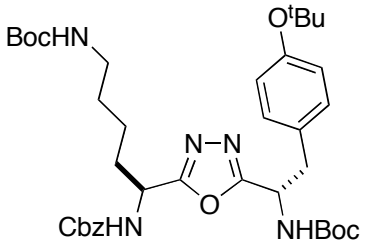
^{1}H NMR (300 MHz, $CDCl_3$) δ 7.43-7.27 (m, 5H), 6.94 (d, 2H, J = 8.2 Hz), 6.86 (d, 2H, J = 8.2 Hz), 5. 94 (br, 1H), 5.38-4.98 (m, 5H), 3.22-2.82 (m, 4H), 1.46-1.20 (m, 18H)

^{13}C NMR (75 MHz, $CDCl_3$) δ 169.3, 166.7, 165.8, 155.6, 154.7, 154.5, 135.9, 129.9, 129.8, 128.6, 128.3, 128.2, 124.2, 82.2, 80.4, 78.4, 67.4, 48.3, 44.5, 39.1, 38.1, 28.8, 28.2, 27.9

MS (ESI, m/z) 639.3 $(M+H)^+$







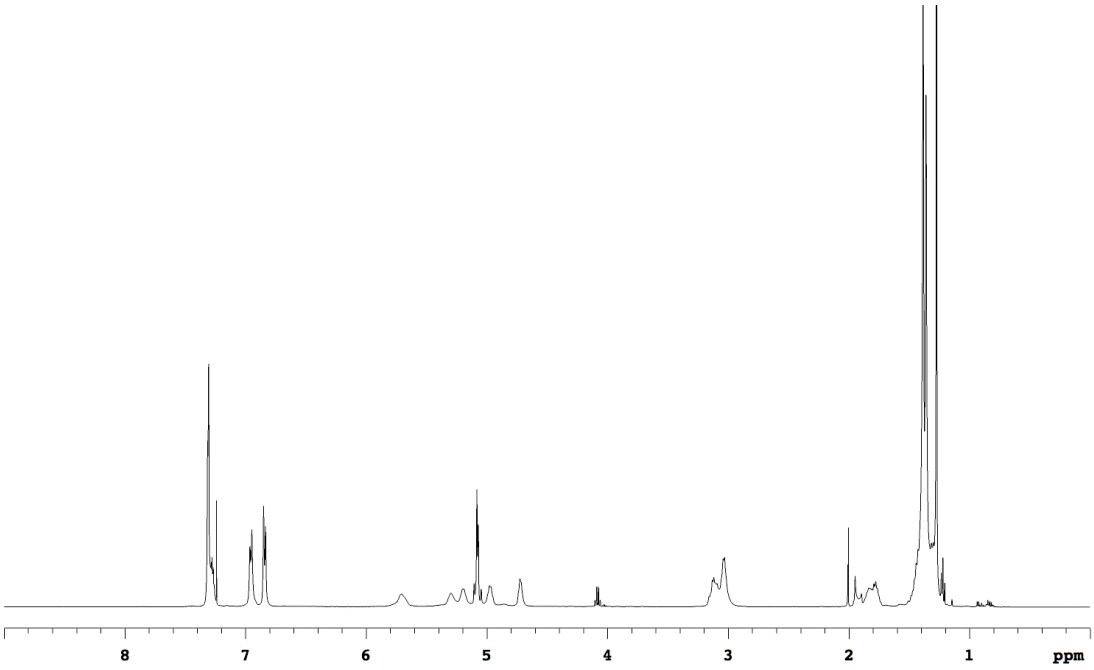
11l

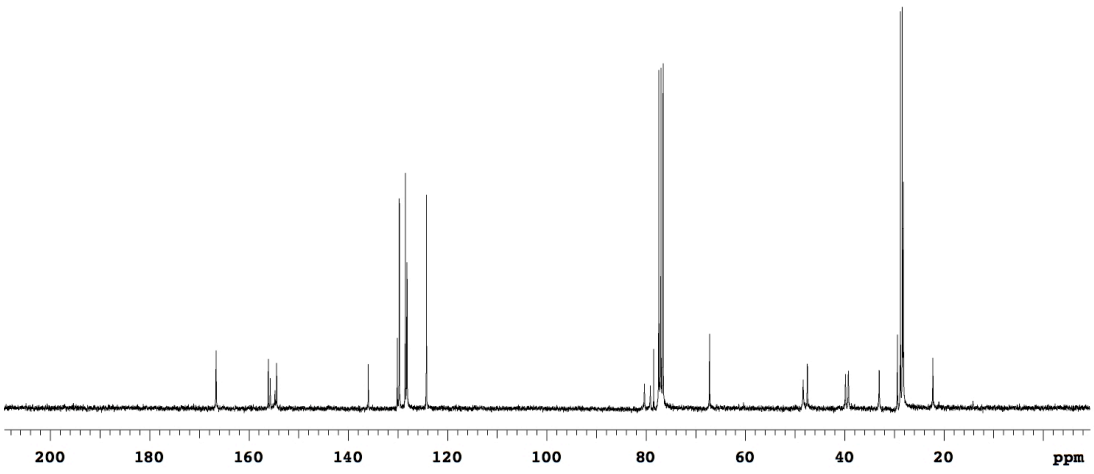
Compound **11l** was prepared from **10d** (4.78 g, 13.6 mmol) and Cbz-Lys(Boc)-OH (5.17 g, 13.6 mmol). Flash chromatography (20 % to 30 % EtOAc/Hexanes) afforded 6.20 g (66 %) **11l** as a white solid.

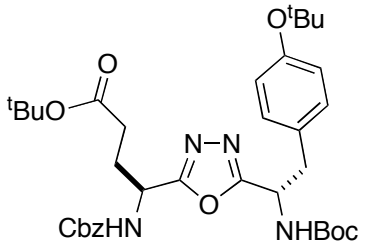
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.34-7.25 (m, 5H), 6.96 (d, 2H, J = 7.9 Hz), 6.84 (d, 2H, J = 7.9 Hz), 5. 71 (br, 1H), 5.36-4.93 (m, 5H), 4.73 (br, 1H), 3.20-2.94 (m, 4H), 1.98-1.72 (m, 2H), 1.53-1.20 (m, 31H)

^{13}C NMR (75 MHz, $CDCl_3$) δ 166.0, 156.1, 155.7, 154.8, 154.4, 135.9, 130.1, 129.7, 128.5, 128.2, 128.1, 124.2, 80.3, 79.1, 78.5, 67.2, 48.4, 47.5, 39.8, 39.2, 33.0, 29.4, 28.7, 28.4, 28.2, 22.2

MS (ESI, m/z) 718.4 $(M+Na)^+$


8
7
6
5
4
3
2
1
ppm


200
180
160
140
120
100
80
60
40
20
ppm



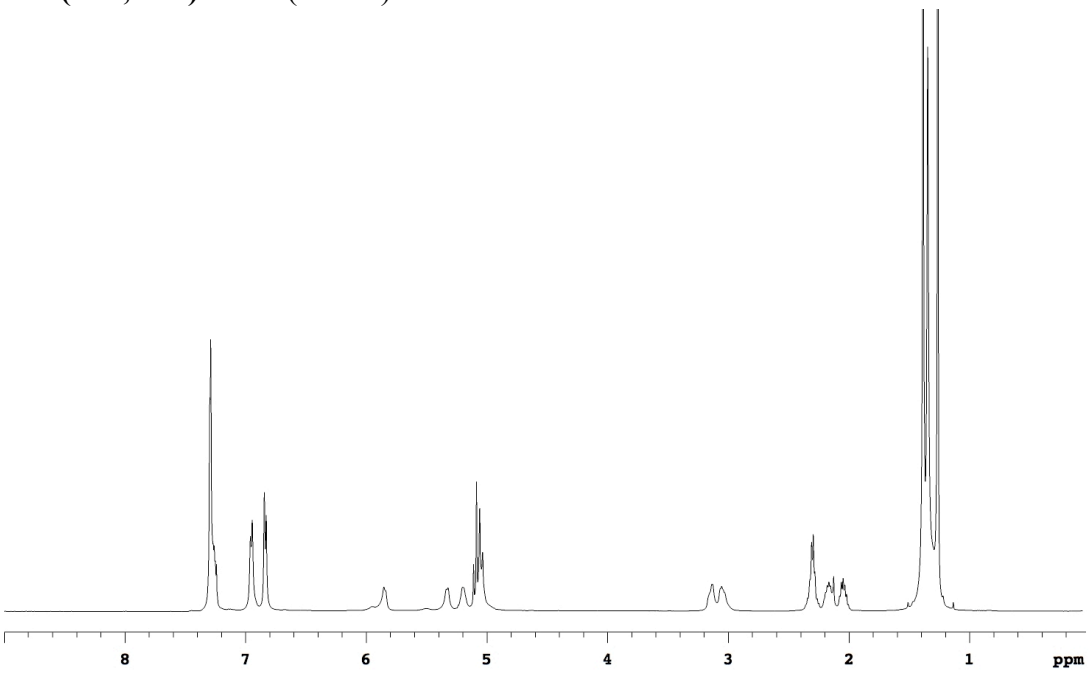
11m

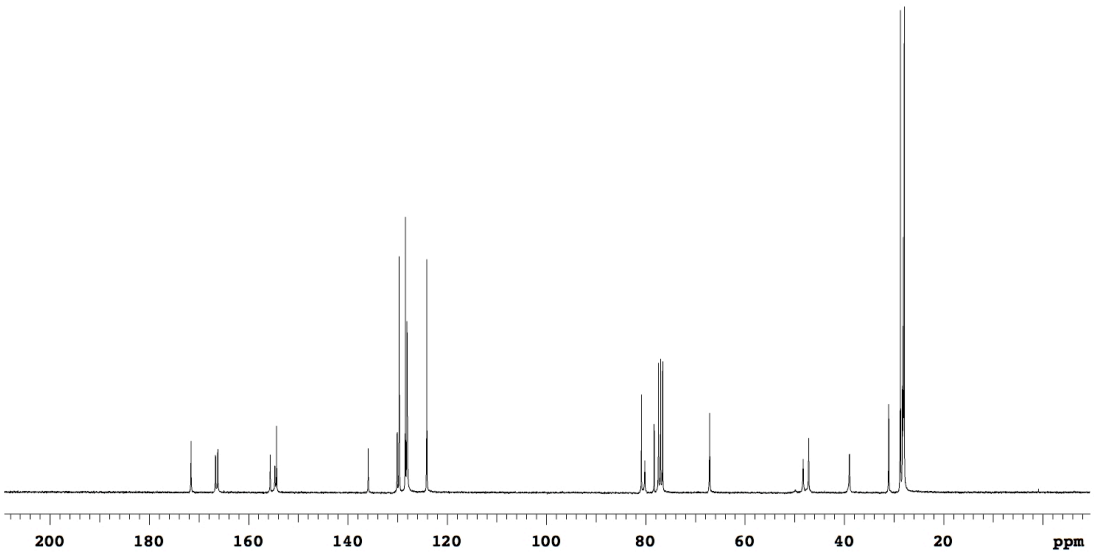
Compound **11m** was prepared from **10d** (5.00 g, 14.2 mmol) and Cbz-Glu(O^{t}Bu)-OH (4.80 g, 14.2 mmol). Flash chromatography (20 % to 30 % EtOAc/Hexanes) afforded 4.79 g (52 %) **11m** as a white solid.

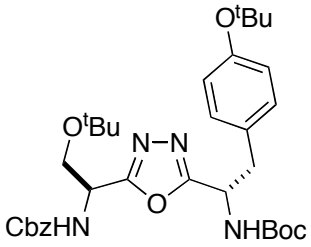
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.35-7.22 (m, 5H), 6.95 (d, 2H, J = 7.8 Hz), 6.84 (d, 2H, J = 7.8 Hz), 5. 85 (br, 1H), 5.32 (br, 1H), 5.20 (br, 1H), 5.14-4.93 (m, 3H), 3.22-2.93 (m, 2H), 2.38-2.24 (m, 2H), 2.23-1.98 (m, 2H), 1.48-1.20 (m, 27H)

^{13}C NMR (75 MHz, $CDCl_3$) δ 171.6, 166.6, 166.2, 155.6, 154.7, 154.3, 135.9, 130.1, 129.6, 128.4, 128.1, 128.0, 124.1, 80.9, 80.2, 78.3, 67.1, 48.3, 47.2, 39.0, 31.0, 28.7, 28.3, 28.1, 27.9

MS (ESI, m/z) 675.3 $(M+Na)^+$







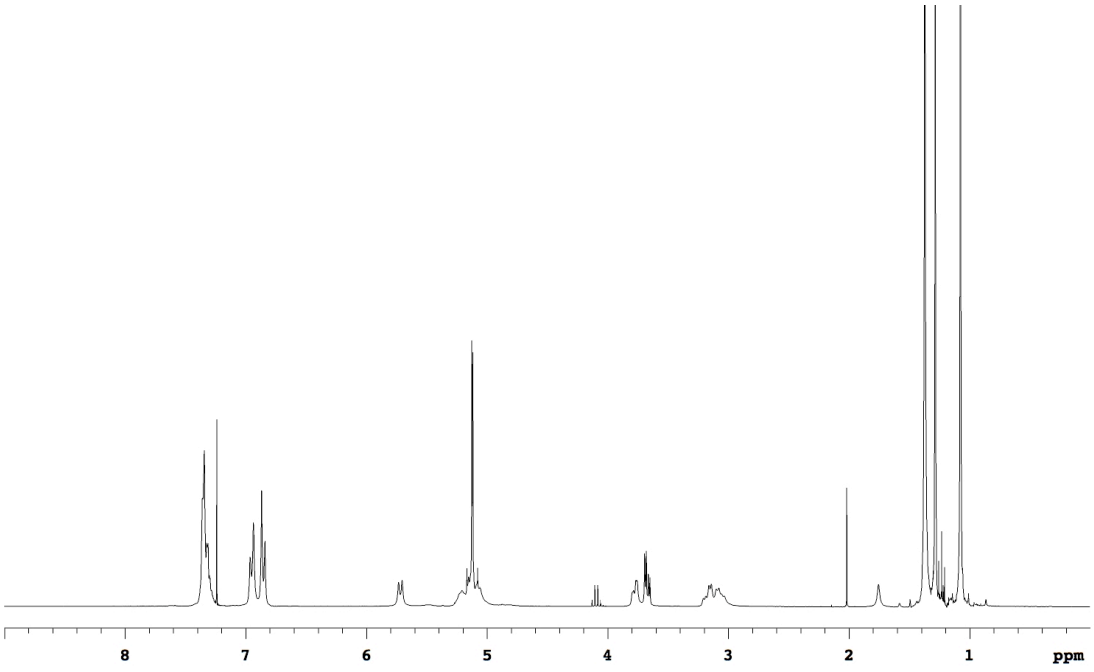
11n

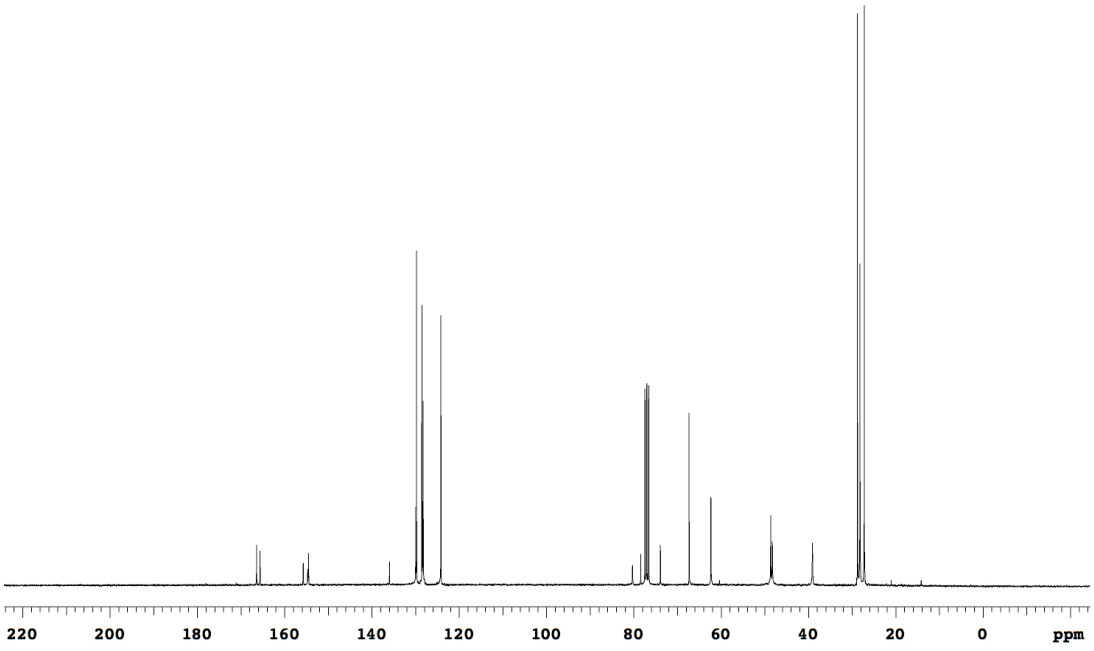
Compound **11n** was prepared from **10d** (5.00 g, 14.2 mmol) and Cbz-Ser(^{t}Bu)-OH (4.20 g, 14.2 mmol). Flash chromatography (20 % to 30 % EtOAc/Hexanes) afforded 3.75 g (43 %) **10d** as a white solid.

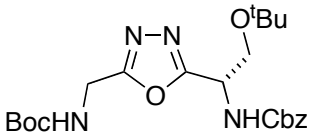
^{1}H NMR (300 MHz, $CDCl_3$) δ 7.40-7.26 (m, 5H), 6.95 (d, 2H, J = 8.3 Hz), 6.85 (d, 2H, J = 8.3 Hz), 5. 72 (d, 1H, J = 8.7 Hz), 5.28-4.98 (m, 5H), 3.78 (dd, 1H, J = 3.0, 9.2 Hz), 3.67 (dd, 1H, J = 3.9, 9.2 Hz), 3.23-2.98 (m, 2H), 1.48-1.19 (m, 18H), 1.08 (s, 9H)

^{13}C NMR (75 MHz, $CDCl_3$) δ 166.4, 165.6, 155.7, 154.7, 154.5, 136.0, 130.0, 129.8, 128.5, 128.3, 128.2, 124.2, 80.4, 78.4, 73.9, 67.3, 62.3, 48.6, 48.3, 39.1, 28.8, 28.2, 27.2

MS (ESI, m/z) 633.3 $(M+Na)^+$


8
7
6
5
4
3
2
1
ppm


220
200
180
160
140
120
100
80
60
40
20
0
ppm



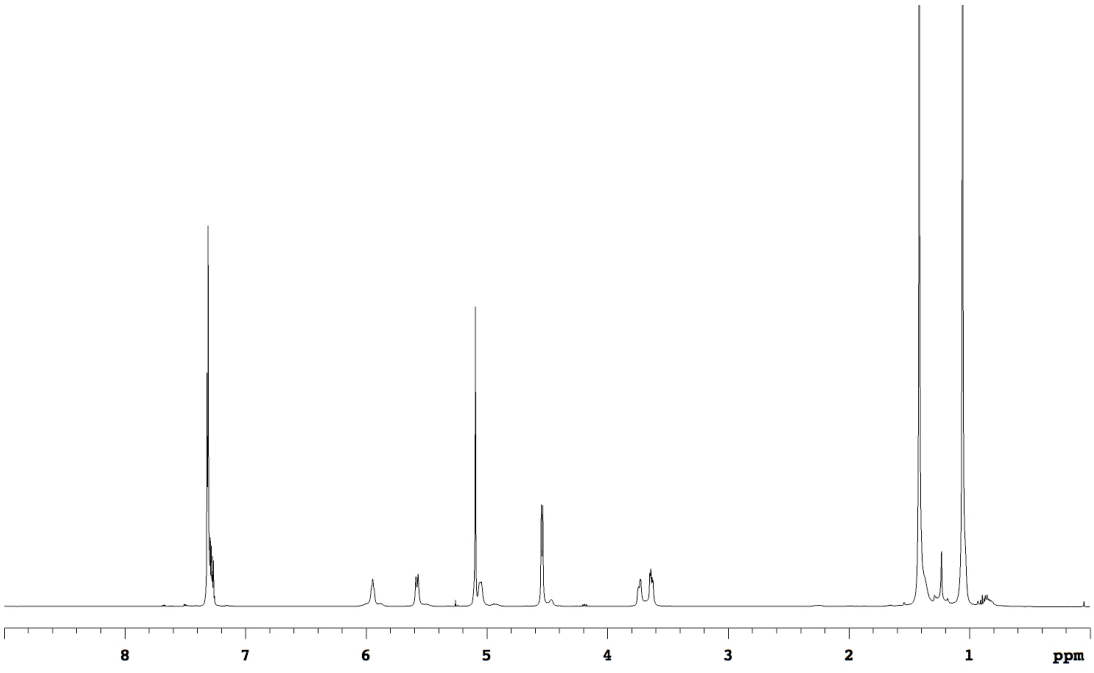
11o

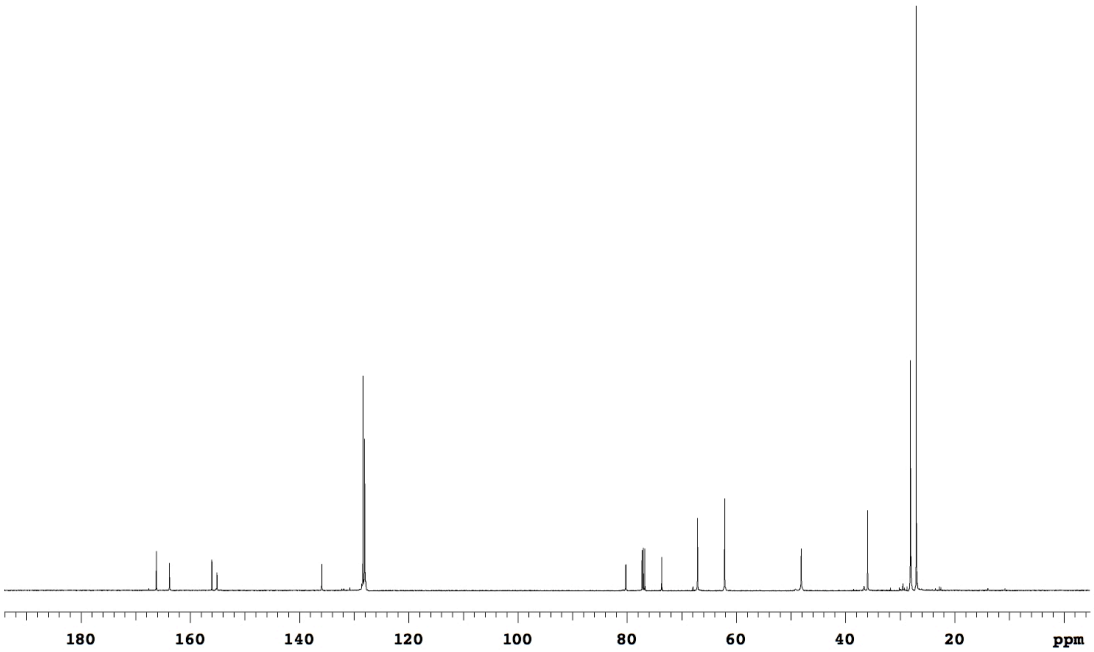
Compound **11o** was prepared from **10a** (0.38 g, 2.0 mmol) and Cbz-Ser(^{t}Bu)-OH (0.59 g, 2.0 mmol). Flash chromatography (20 % to 30 % EtOAc/Hexanes) afforded 0.36 g (40 %) **11o** as a white solid.

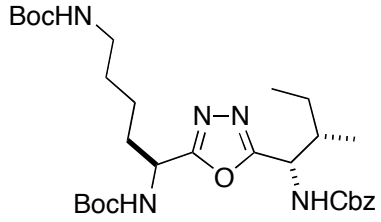
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.36-7.24 (m, 5H), 5.95 (br, 1H), 5.58 (d, 1H, J = 8.8 Hz), 5.14-5.01 (m, 3H), 4.54 (d, 2H, J = 6.0 Hz), 3.77-3.70 (m, 1H), 3.64 (dd, 1H, J = 4.0, 9.1 Hz), 1.42 (s, 9H), 1.06 (s, 9H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 166.2, 163.8, 156.0, 155.1, 135.9, 128.4, 128.1, 128.0, 80.2, 73.7, 67.1, 62.1, 48.1, 36.0, 28.1, 27.0

MS (ESI, m/z) 471.2 $(M+Na)^+$







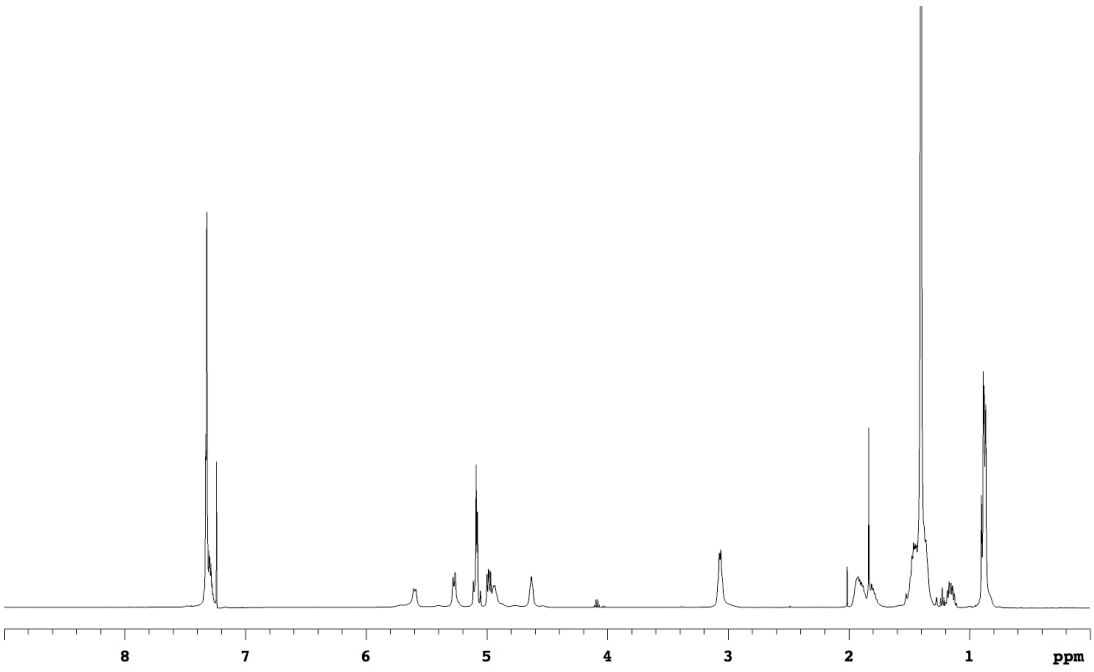
11p

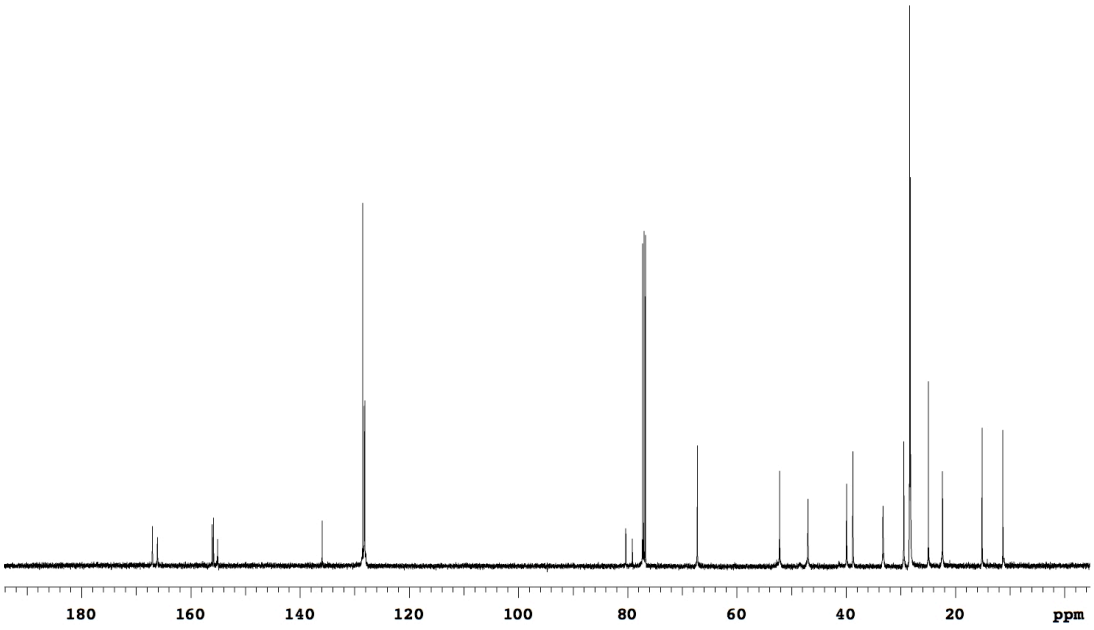
Compound **11p** was prepared from **10c** (2.26 g, 6.3 mmol) and Cbz-Ile-OH•DCHA (2.80 g, 6.3 mmol). Flash chromatography (20 % to 30 % EtOAc/Hexanes) afforded 2.26 g (61 %) **11p** as a white solid.

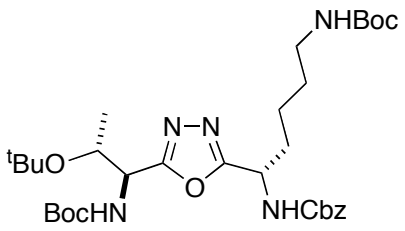
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.39-7.26 (m, 5H), 5.60 (d, 1H, J = 8.5 Hz), 5.27 (d, 1H, J = 8.6 Hz), 5.14-5.03 (m, 2H), 5.02-4.84 (m, 2H), 4.63 (br, 1H), 3.13-3.00 (m, 2H), 2.00-1.73 (m, 3H), 1.55-1.30 (m, 23H), 1.21-1.10 (m, 1H), 0.94-0.80 (m, 6H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 167.0, 166.1, 156.1, 155.8, 155.1, 135.9, 128.5, 128.2, 128.1, 80.3, 79.1, 67.2, 52.2, 47.0, 39.9, 38.8, 33.2, 29.4, 28.4, 28.2, 24.9, 22.3, 15.1, 11.3

MS (ESI, m/z) 590.4 $(M+H)^+$


8
7
6
5
4
3
2
1
ppm


180
160
140
120
100
80
60
40
20
ppm



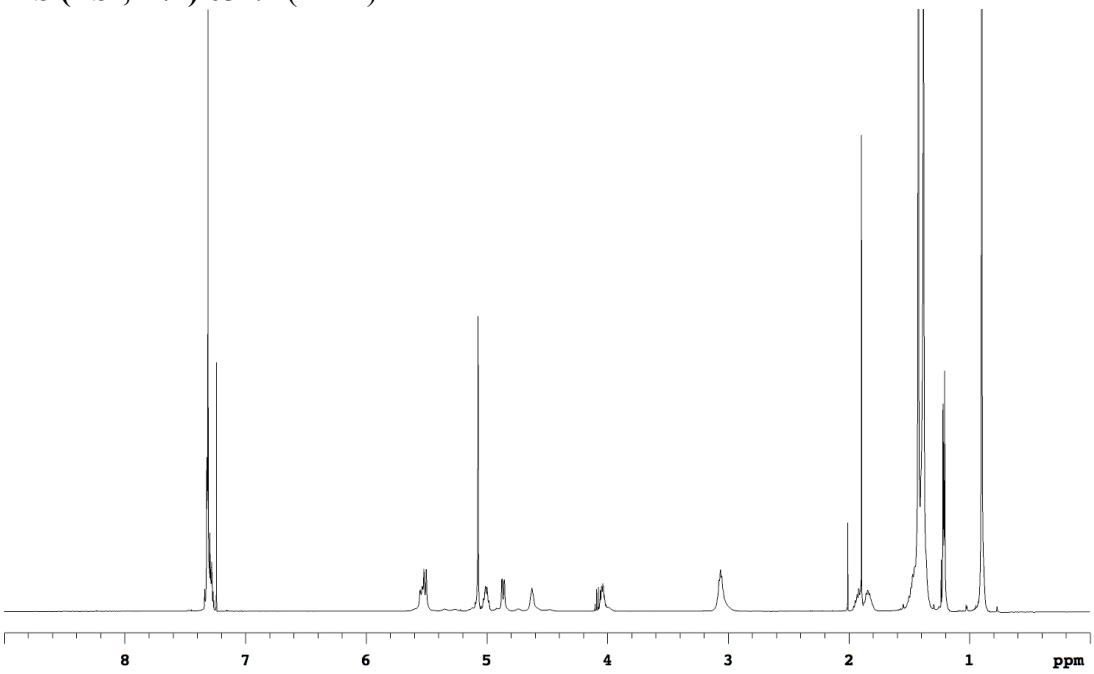
11q

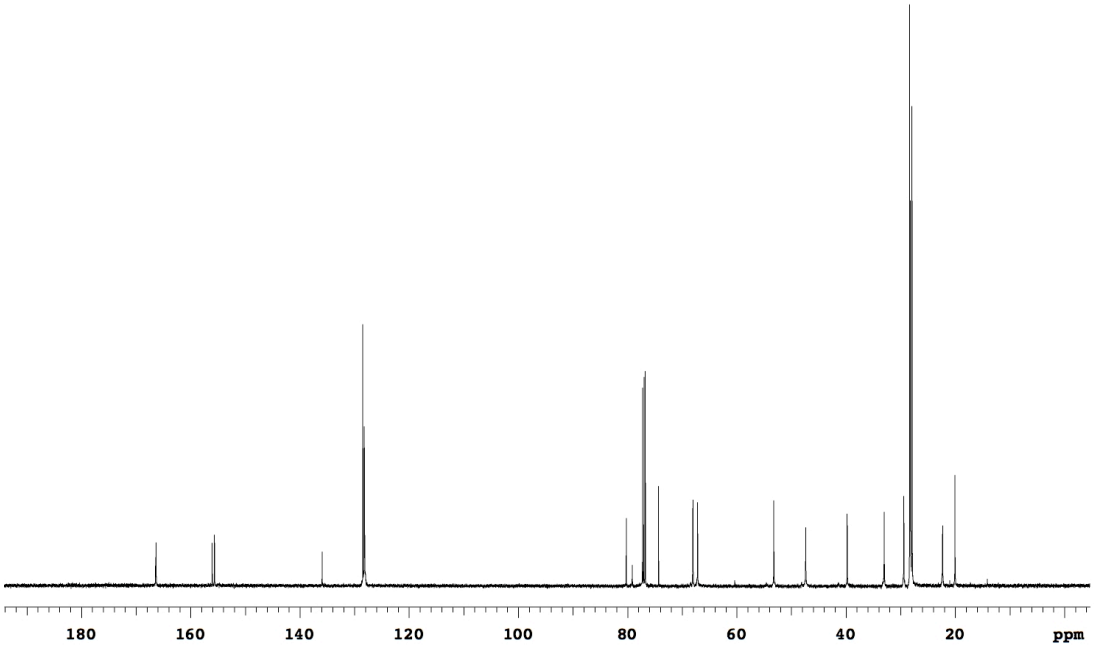
Compound **11q** was prepared from **10f** (2.59 g, 8.9 mmol) and Cbz-Lys(Boc)-OH (3.48 g, 8.9 mmol). Flash chromatography (20 % to 30 % EtOAc/Hexanes) afforded 3.42 g (60%) **11q** as a white solid.

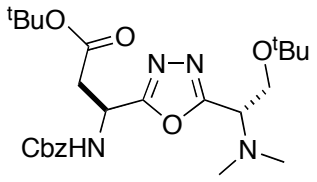
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.37-7.26 (m, 5H), 5.60-5.46 (m, 2H), 5.07 (s, 2H), 5.05-4.97 (m, 1H), 4.87 (dd, 1H, J = 1.5, 9.6 Hz), 4.63 (br, 1H), 4.10-3.95 (m, 1H), 3.14-2.95 (m, 2H), 1.98-1.77 (m, 2H), 1.53-1.30 (m, 22H), 1.21 (d, 3H, J = 6.1 Hz), 0.90 (s, 9H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 166.4, 166.3, 156.1, 155.7, 155.1, 135.9, 128.5, 128.2, 128.1, 80.2, 79.1, 74.3, 68.0, 67.2, 53.2, 47.4, 39.8, 33.0, 29.4, 28.3, 28.2, 27.9, 22.3, 20.0

MS (ESI, m/z) 634.4 $(M+H)^+$







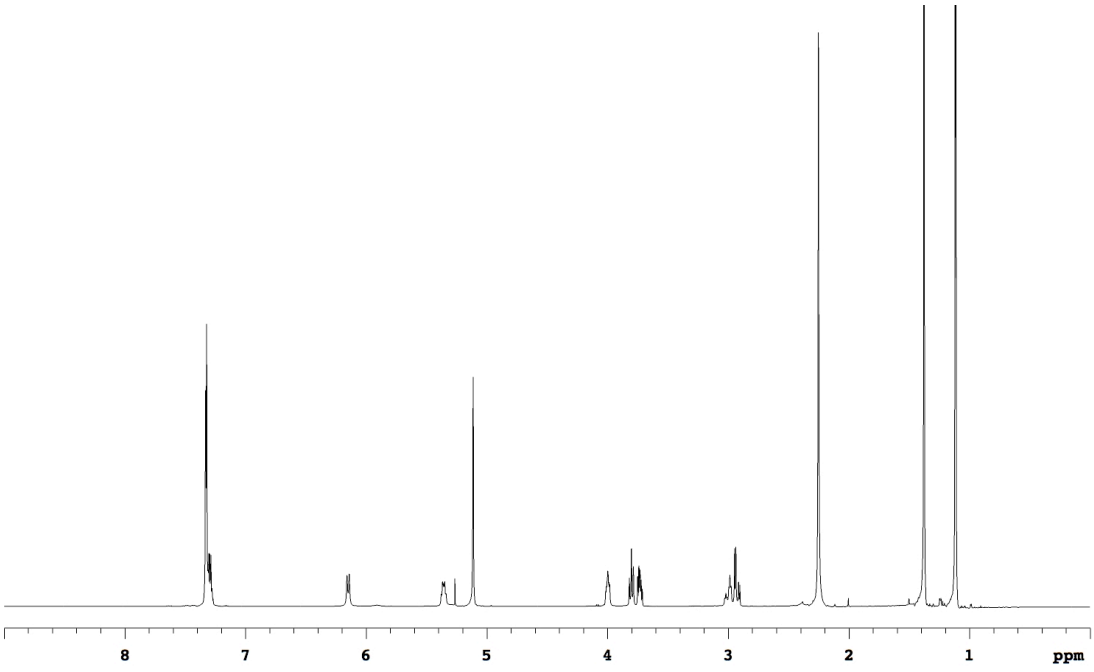
11r

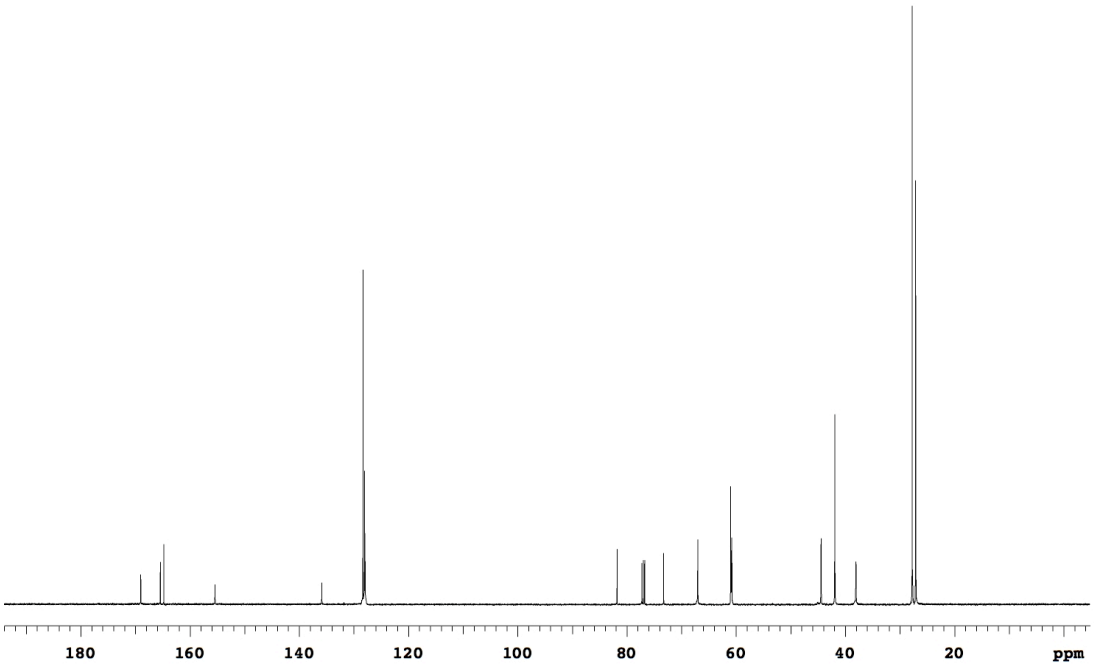
Compound **11r** was prepared from **10g** (1.42 g, 7.0 mmol) and Cbz-Asp(O^{t}Bu)-OH•DCHA (3.53 g, 7.0 mmol). Flash chromatography (20 % to 50 % EtOAc/Hexanes) afforded 1.51 g (44 %) **11r** as a white solid.

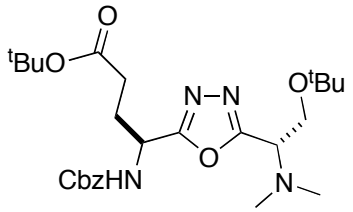
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.38-7.26 (m, 5H), 6.15 (d, 1H, J = 9.2 Hz), 5.36 (dt, 1H, J = 5.2, 9.2 Hz), 5.11 (s, 2H), 4.03-3.96 (m, 1H), 3.84-3.70 (m, 2H), 3.00 (dt, 1H, J = 4.7, 16.6 Hz), 2.92 (dd, 1H, J = 5.4, 16.6 Hz), 2.25 (s, 6H), 1.38 (s, 9H), 1.11 (s, 9H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 169.0, 165.4, 164.8, 155.4, 135.9, 128.3, 128.1, 127.9, 81.8, 73.3, 67.0, 61.0, 60.8, 44.4, 41.9, 38.1, 27.8, 27.1

MS (ESI, m/z) 491.3 $(M+H)^+$


8
7
6
5
4
3
2
1
ppm


180
160
140
120
100
80
60
40
20
ppm



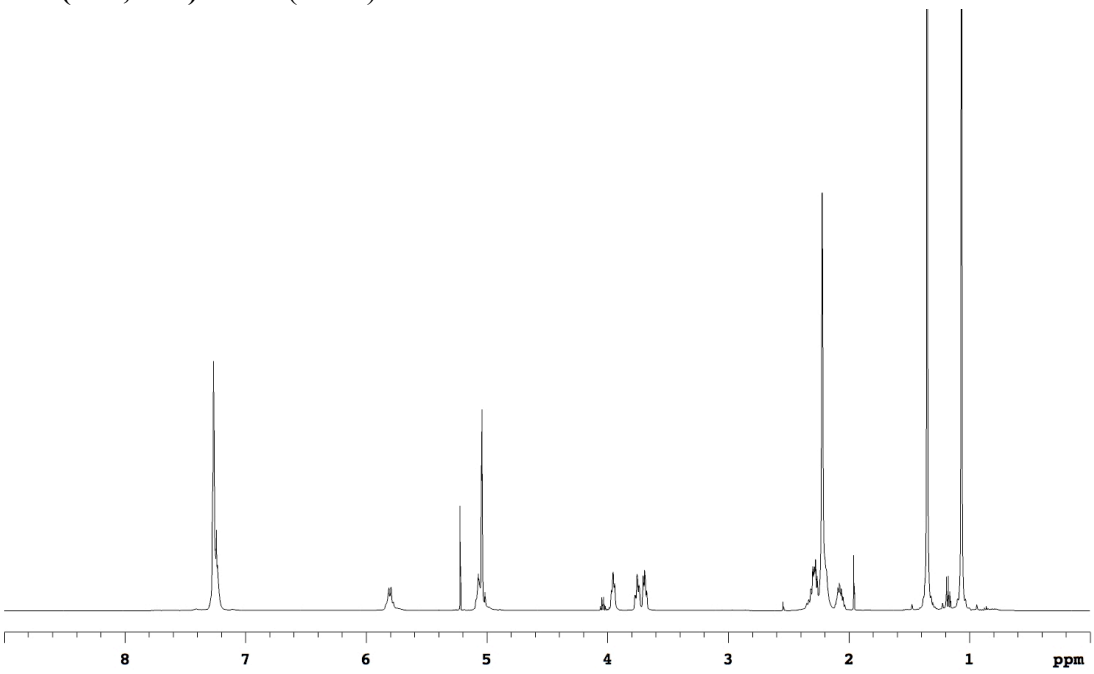
11s

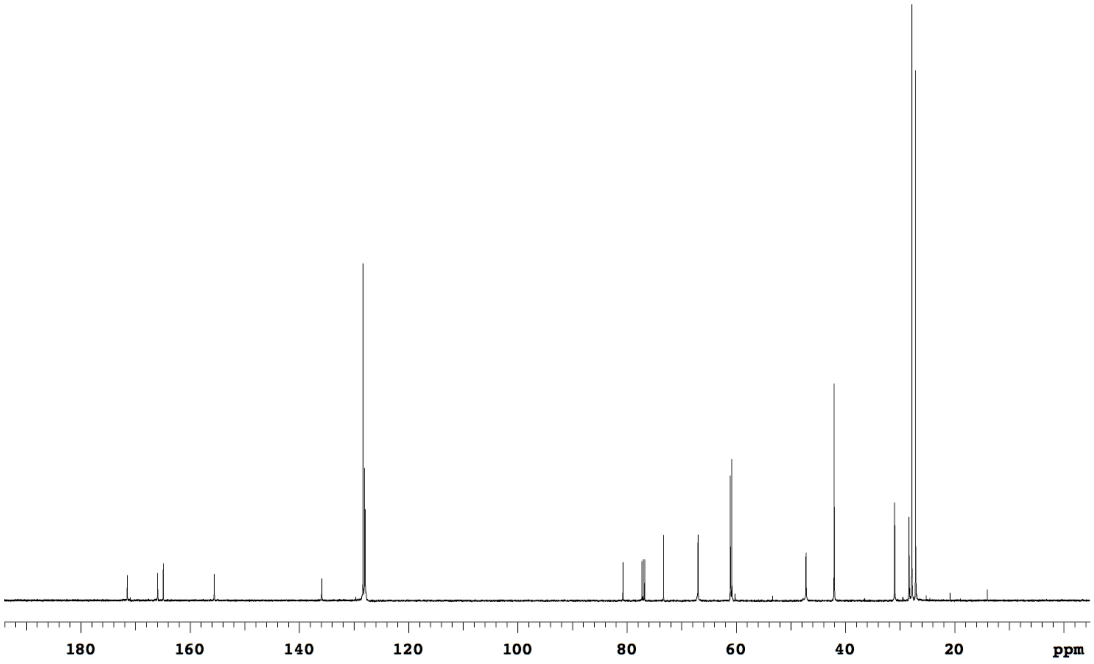
Compound **11s** was prepared from **10g** (1.02 g, 5.0 mmol) and Cbz-Glu(O^{t}Bu)-OH (1.69 g, 7.0 mmol). Flash chromatography (20 % to 50 % EtOAc/Hexanes) afforded 1.02 g (40 %) **11s** as a white solid.

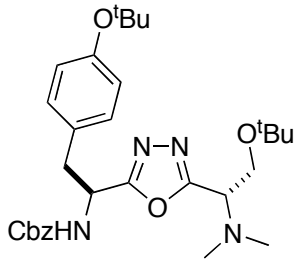
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.33-7.17 (m, 5H), 5.86-5.75 (m, 1H), 5.12-4.97 (m, 3H), 3.99-3.91 (m, 1H), 3.79-3.65 (m, 2H), 2.36-2.14 (m, 9H), 2.12-2.02 (m, 1H), 1.35 (s, 9H), 1.07 (s, 9H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 171.5, 165.9, 164.9, 155.6, 135.9, 128.3, 128.0, 127.9, 80.7, 73.3, 67.0, 61.1, 60.8, 47.2, 42.1, 31.0, 28.4, 27.8, 27.2

MS (ESI, m/z) 505.3 $(M+H)^+$







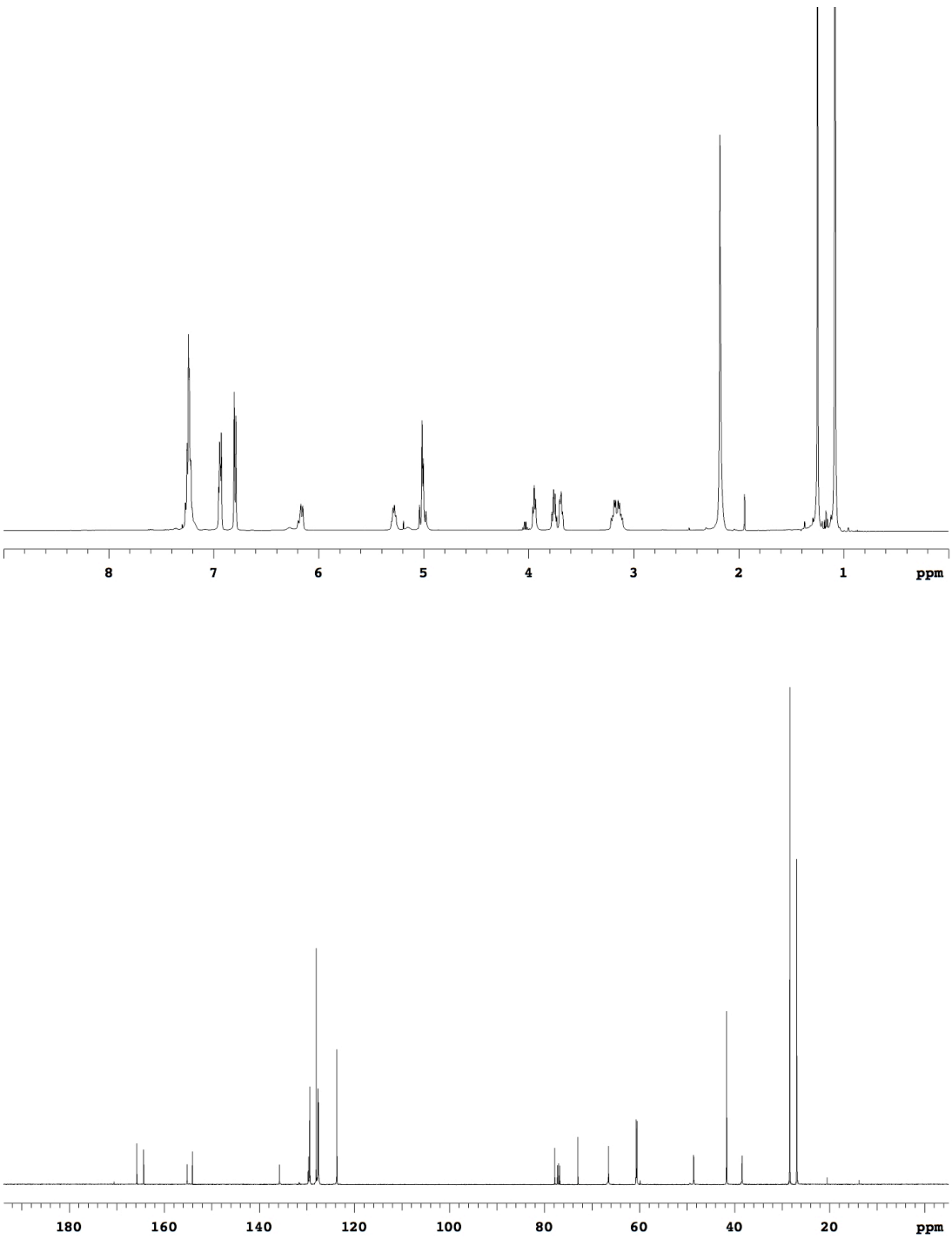
11t

Compound **11t** was prepared from **10g** (2.66 g, 7.0 mmol) and Cbz-Tyr(^{t}Bu)-OH (2.60 g, 7.0 mmol). Flash chromatography (20 % to 50 % EtOAc/Hexanes) afforded 2.72 g (72 %) **11t** as a white solid.

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.30-7.14 (m, 5H), 6.98-6.90 (m, 2H), 6.80 (d, 2H, J = 8.0 Hz), 6.22-6.13 (m, 1H), 5.33-5.23 (m, 1H), 5.06-4.94 (m, 2H), 3.99-3.90 (m, 1H), 3.80-3.66 (m, 2H), 3.26-3.07 (m, 2H), 2.18 (s, 6H), 1.25 (s, 9H), 1.09 (s, 9H)

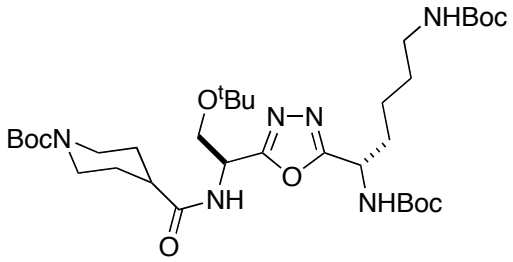
^{13}C NMR (125 MHz, $CDCl_3$) δ 165.8, 164.3, 155.2, 154.1, 135.8, 129.6, 129.4, 128.0, 127.7, 127.6, 123.7, 77.8, 73.0, 66.5, 60.7, 60.5, 48.6, 41.7, 38.4, 28.4, 26.9

MS (ESI, m/z) 539.3 $(M+H)^+$



General Procedure for Preparation of 12a-t. A solution of Cbz protected oxadizole **11a-t** in dry methanol (0.10 M) was stirred vigorously with 10 % Pd-C (0.1

equiv) in a H_2 atmosphere for 12 h at 25 °C until the reaction is complete. Filtration and evaporation of the solvent *in vacuo* yielded the crude product and it was used directly in the next step. The amine residue (1.0 equiv) and Boc-Inp-OH (1.0 equiv) was dissolved in dry CH_2Cl_2 (0.15 M). *N*-methylmorpholine (2.1 equiv) was added to the above solution followed with HOBt (1.1 equiv) and EDC (1.1 equiv). The resulting solution was stirred at room temperature for 12h. The reaction mixture was washed with H_2O (30 ml), 1N HCl (30 ml), saturated $NaHCO_3$ (30 ml) and brine (30 ml). The organic layer was separated and dried with Na_2SO_4 and concentrated to dryness to afford the crude product. Flash chromatography with EtOAc/hexanes mixtures provided pure compounds **12a-t**.



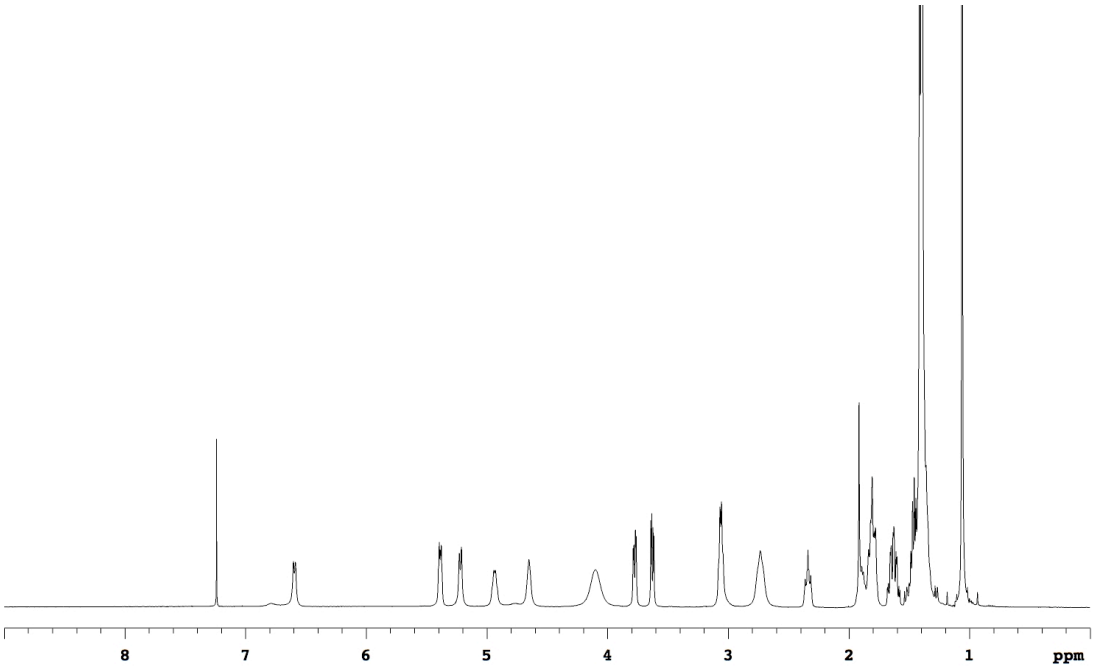
12a

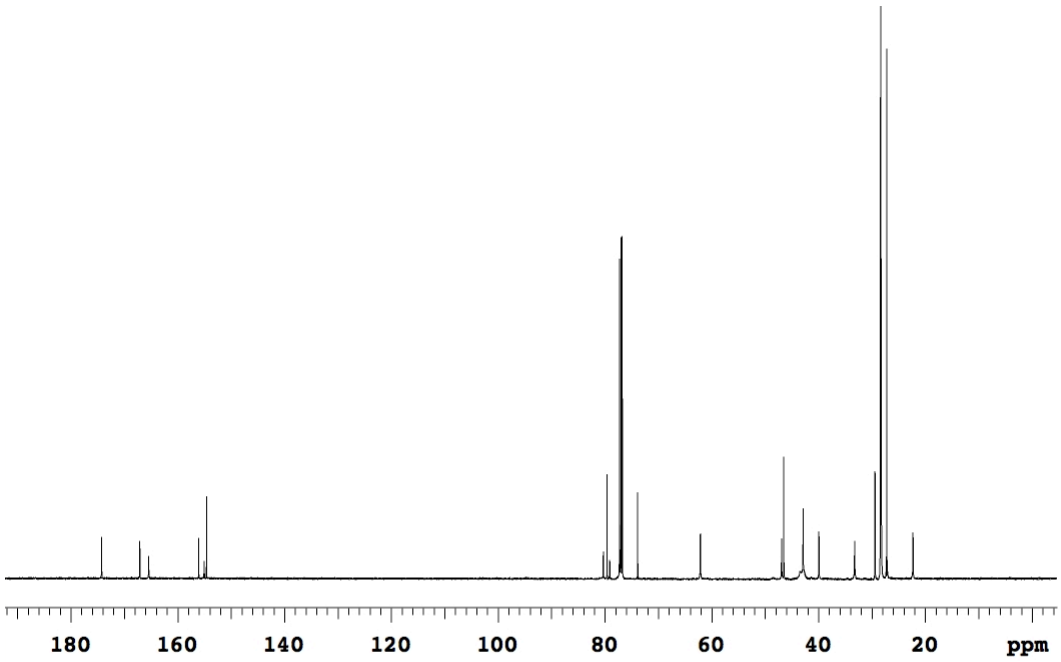
Compound **12a** was prepared from (4.83 g, 7.8 mmol) of **11a**. Flash chromatography (30 % to 50 % EtOAc/Hexanes) afforded 3.78 g (70 %) **12a** as a white solid.

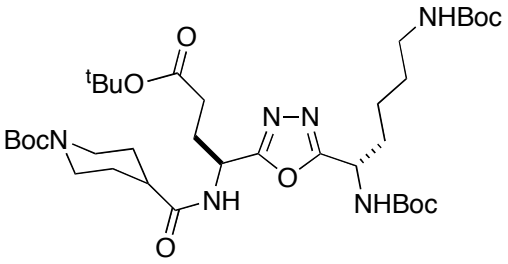
^{1}H NMR (500 MHz, $CDCl_3$) δ 6.59 (d, 1H, J = 8.2 Hz), 5.42-5.36 (m, 1H), 5.22 (d, 1H, J = 8.5 Hz), 4.98-4.89 (m, 1H), 4.65 (br, 1H), 4.10 (br, 2H), 3.78 (dd, 1H, J = 3.2, 9.4 Hz), 3.15-2.94 (dd, 1H, J = 3.9, 9.4 Hz), 3.12-2.99 (m, 2H), 2.81-2.65 (br, 2H), 2.39-2.30 (m, 1H), 1.94-1.75 (m, 4H), 1.70-1.57 (m, 2H), 1.51-1.31 (m, 31H), 1.06 (s, 9H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 174.2, 167.1, 165.4, 156.1, 155.0, 154.6, 80.3, 80.0, 79.6, 79.1, 73.8, 62.1, 46.9, 46.5, 42.9, 39.9, 33.2, 29.4, 28.4, 28.4, 28.3, 28.2, 27.2, 22.3

MS (ESI, m/z) 719.4 $(M+Na)^+$


8
7
6
5
4
3
2
1
ppm


180
160
140
120
100
80
60
40
20
ppm



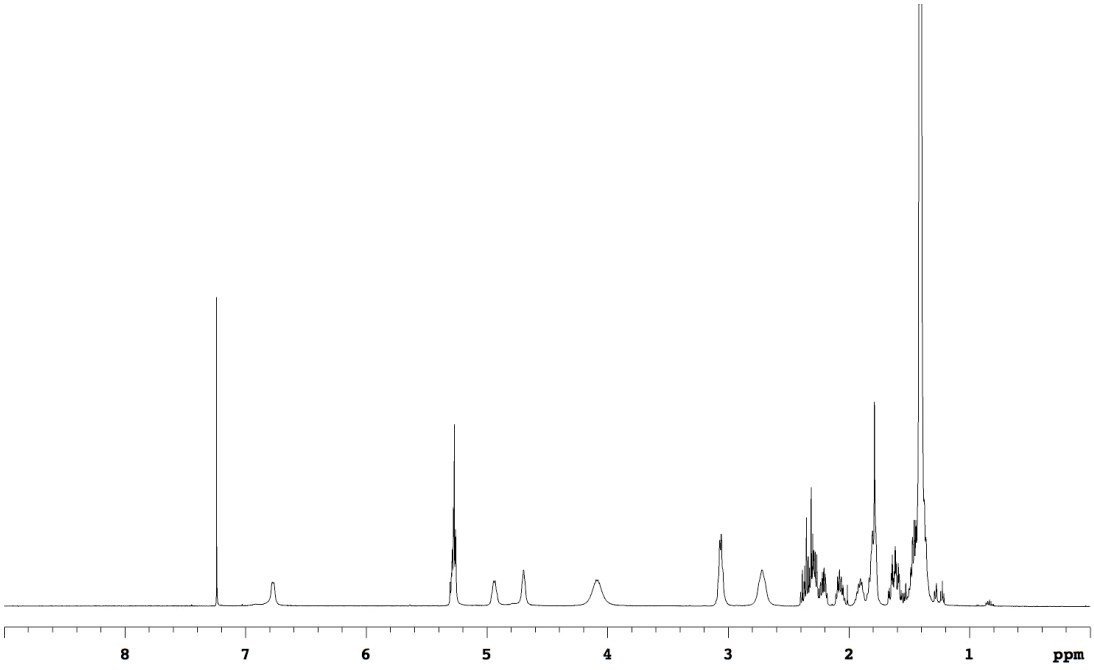
12b

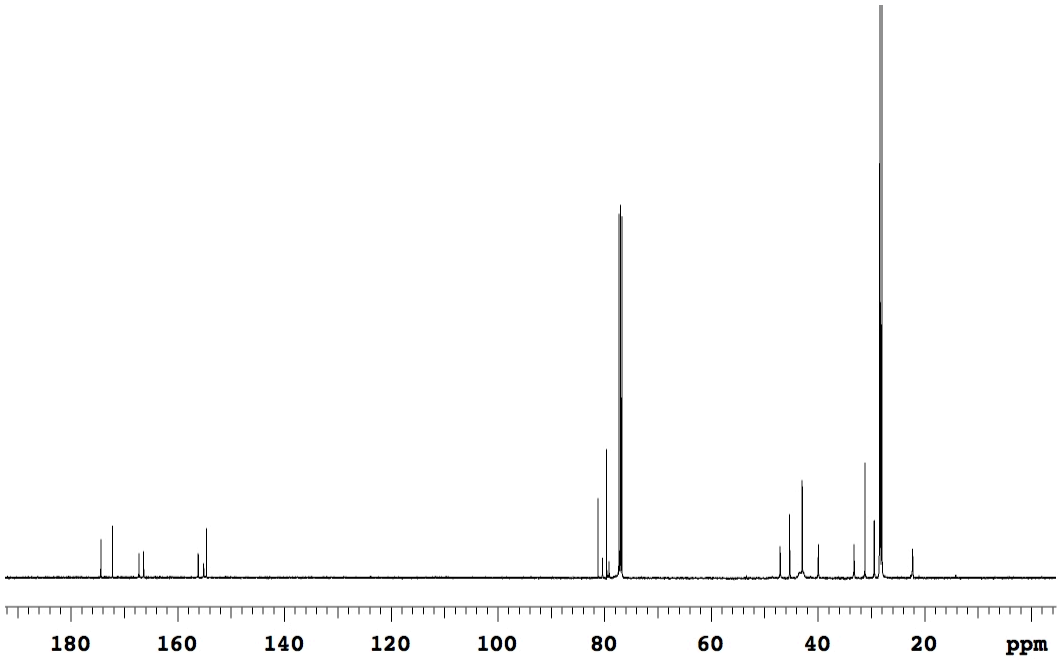
Compound **12b** was prepared from (4.43 g, 6.7 mmol) of **11b**. Flash chromatography (30 % to 50 % EtOAc/Hexanes) afforded 3.83 g (77 %) **12b** as a white solid.

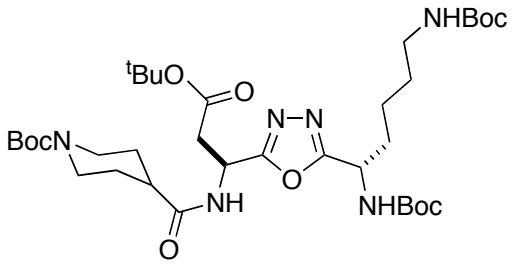
^{1}H NMR (500 MHz, $CDCl_3$) δ 6.77 (br, 1H), 5.32-5.23 (m, 2H), 4.98-4.89 (m, 1H), 4.70 (br, 1H), 4.09 (br, 2H), 3.12-3.01 (m, 2H), 2.72 (br, 2H), 2.41-2.17 (m, 4H), 2.12-2.02 (m, 1H), 1.96-1.74 (m, 4H), 1.69-1.55 (m, 2H), 1.52-1.32 (m, 40H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 174.4, 172.2, 167.3, 166.4, 156.1, 155.1, 154.6, 81.2, 80.4, 79.6, 79.1, 47.1, 45.3, 42.9, 39.9, 33.2, 31.2, 29.4, 28.5, 28.4, 28.4, 28.3, 28.3, 28.0, 22.3

MS (MALDI, m/z) 761.1 $(M+Na)^+$







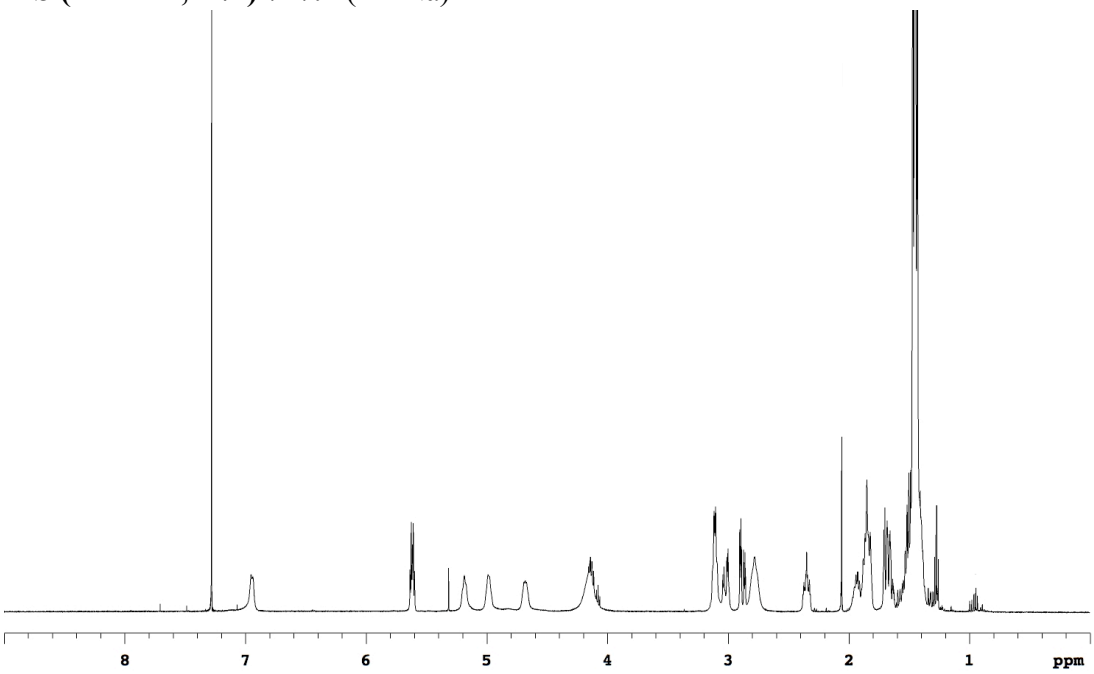
12c

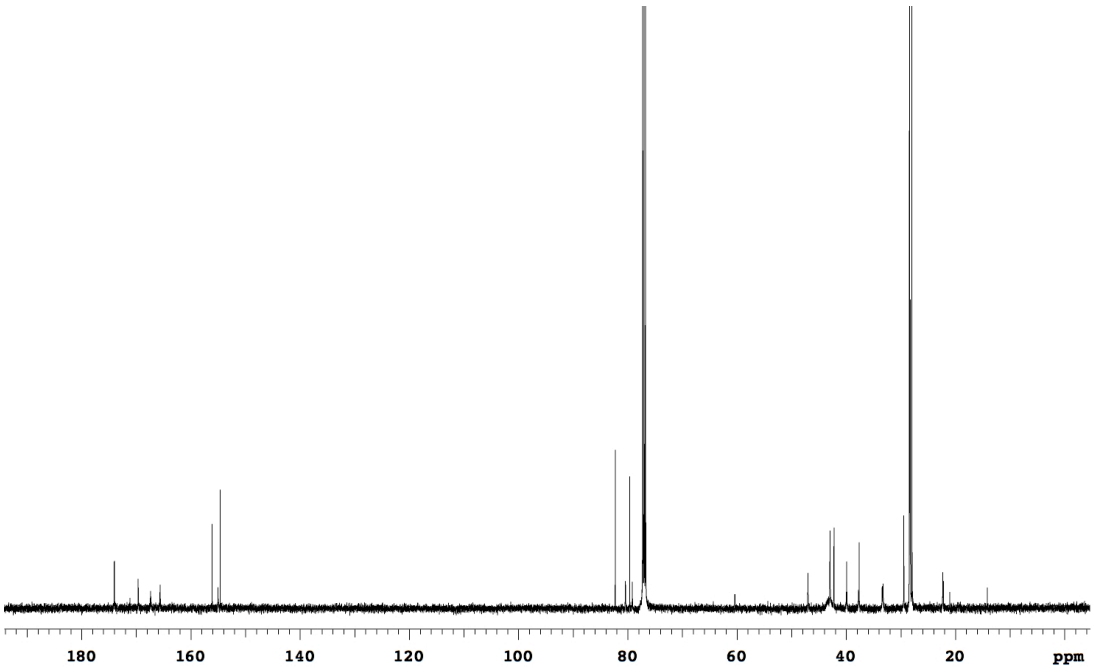
Compound **12c** was prepared from (3.27 g, 5.0 mmol) of **11c**. Flash chromatography (30 % to 50 % EtOAc/Hexanes) afforded 2.60 g (71 %) **12c** as a white solid.

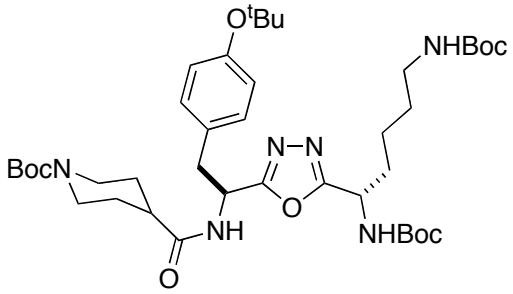
^{1}H NMR (500 MHz, $CDCl_3$) δ 6.95 (br, 1H), 5.66-5.58 (m, 1H), 5.24-5.13 (m, 1H), 5.03-4.93 (m, 1H), 4.73-4.63 (m, 1H), 4.25-4.05 (br, 2H), 3.16-3.06 (m, 2H), 3.06-2.97 (m, 1H), 2.88 (ddd, 1H, J = 2.3, 5.2, 16.5 Hz), 2.78 (br, 2H), 2.35 (tt, 1H, J = 3.7, 11.5 Hz), 2.00-1.79 (m, 4H), 1.73-1.62 (m, 2H), 1.57-1.36 (m, 40H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 174.0, 169.6, 167.4, 165.6, 156.1, 155.0, 154.6, 82.3, 80.4, 79.6, 79.2, 47.0, 43.0, 42.2, 40.0, 37.6, 33.4, 33.2, 29.4, 28.4, 28.3, 28.3, 28.0, 22.3, 22.2

MS (MALDI, m/z) 747.2 $(M+Na)^+$







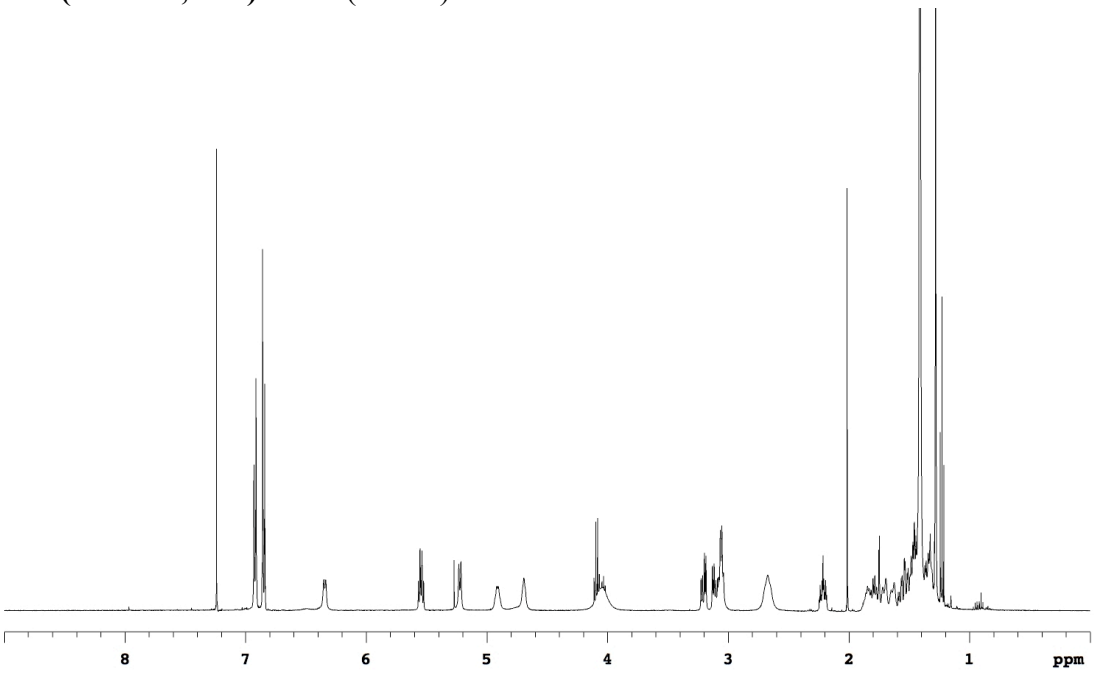
12d

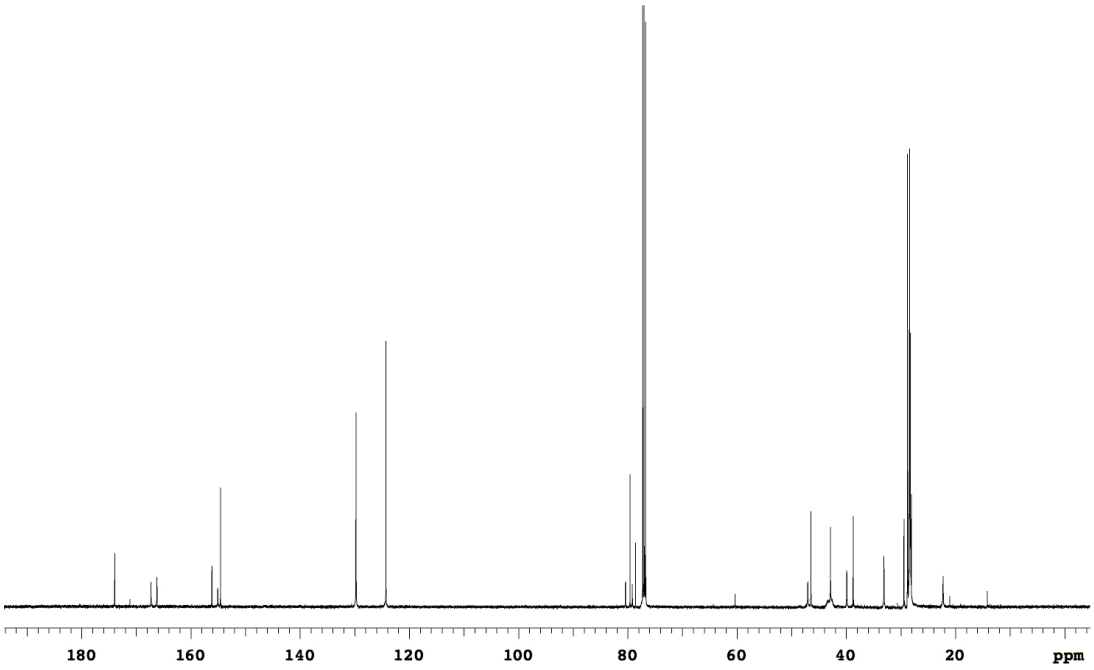
Compound **12d** was prepared from (3.66 g, 5.3 mmol) of **11d**. Flash chromatography (30 % to 50 % EtOAc/Hexanes) afforded 3.80 g (93 %) **12d** as a white solid.

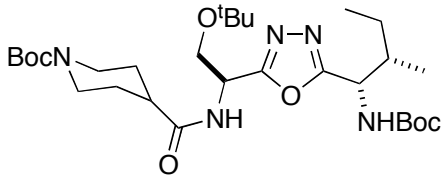
^{1}H NMR (500 MHz, $CDCl_3$) δ 6.92 (d, 2H, J = 8.5 Hz), 6.85 (d, 2H, J = 8.5 Hz), 6.34 (d, 1H, J = 8.2 Hz), 5.58-5.51 (m, 1H), 5.22 (d, 1H, J = 8.5 Hz), 4.96- 4.86 (m, 1H), 4.69 (br, 1H), 4.05 (br, 2H), 3.21 (dd, 1H, J = 6.2, 14.0 Hz), 3.15-3.01 (m, 3H), 2.74-2.60 (br, 2H), 2.21 (tt, 1H, J = 3.7, 11.5 Hz), 1.91-1.60 (m, 4H), 1.60-1.30 (m, 33H), 1.28 (s, 9H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 174.0, 171.2, 167.3, 166.2, 156.1, 155.1, 154.6, 129.8, 129.7, 124.3, 80.4, 79.2, 78.6, 47.0, 45.5, 42.9, 39.9, 38.7, 33.1, 29.4, 28.8, 28.6, 28.4, 28.4, 28.3, 28.1, 22.3

MS (MALDI, m/z) 795.2 $(M+Na)^+$







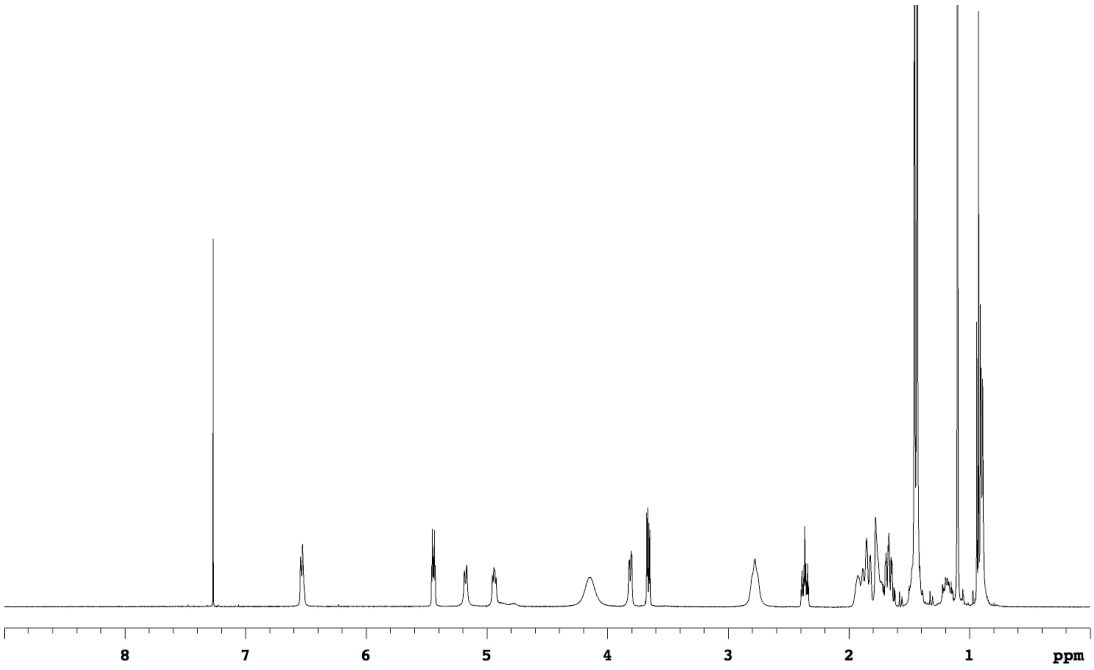
12e

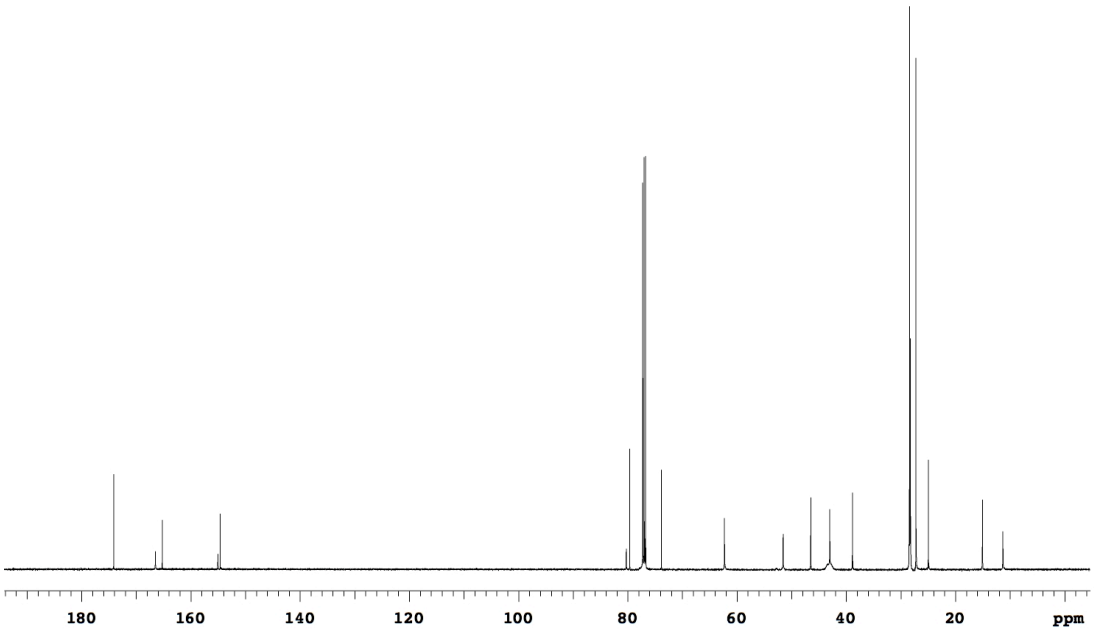
Compound **12e** was prepared from (5.26 g, 10.4 mmol) of **11e**. Flash chromatography (30 % to 50 % EtOAc/Hexanes) afforded 4.80 g (78 %) **12e** as a white solid.

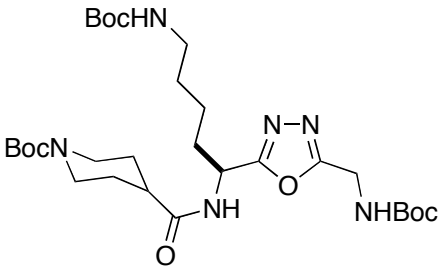
^{1}H NMR (500 MHz, $CDCl_3$) δ 6.53 (d, 1H, J = 8.6 Hz), 5.44 (dt, 1H, J = 3.5, 8.6 Hz), 5.18 (d, 1H, J = 9.2 Hz), 4.98-4.90 (m, 1H), 4.14 (br, 2H), 3.81 (dd, 1H, J = 2.6, 9.2 Hz), 3.66 (dd, 1H, J = 3.7, 9.2 Hz), 2.78 (br, 2H), 2.37 (tt, 1H, J = 3.7, 11.5 Hz), 1.97-1.61 (m, 5H), 1.51-1.38 (m, 19H), 1.24-1.13 (m, 1H), 1.10 (s, 9H), 0.95-0.87 (m, 6H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 174.1, 166.5, 165.2, 155.1, 154.6, 80.3, 79.6, 73.8, 62.3, 51.6, 46.5, 43.0, 38.9, 28.5, 28.4, 28.3, 28.2, 27.2, 25.0, 15.1, 11.3

MS (MALDI, m/z) 604.1 $(M+Na)^+$


8
7
6
5
4
3
2
1
ppm


180
160
140
120
100
80
60
40
20
ppm



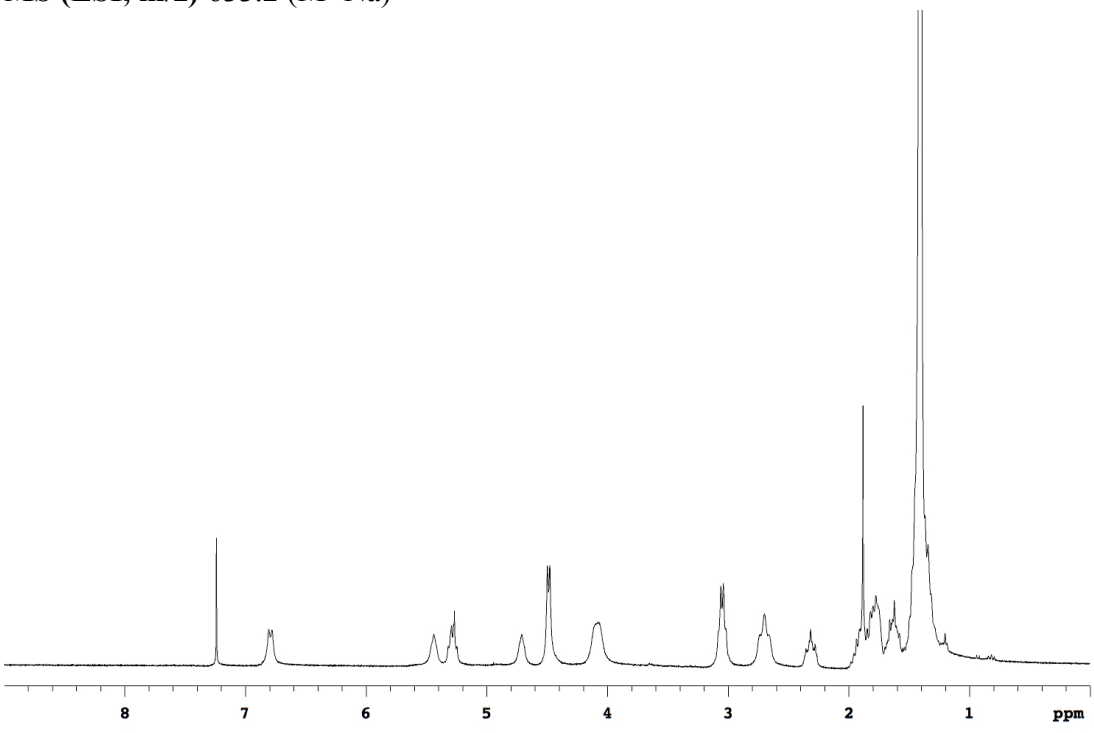
12f

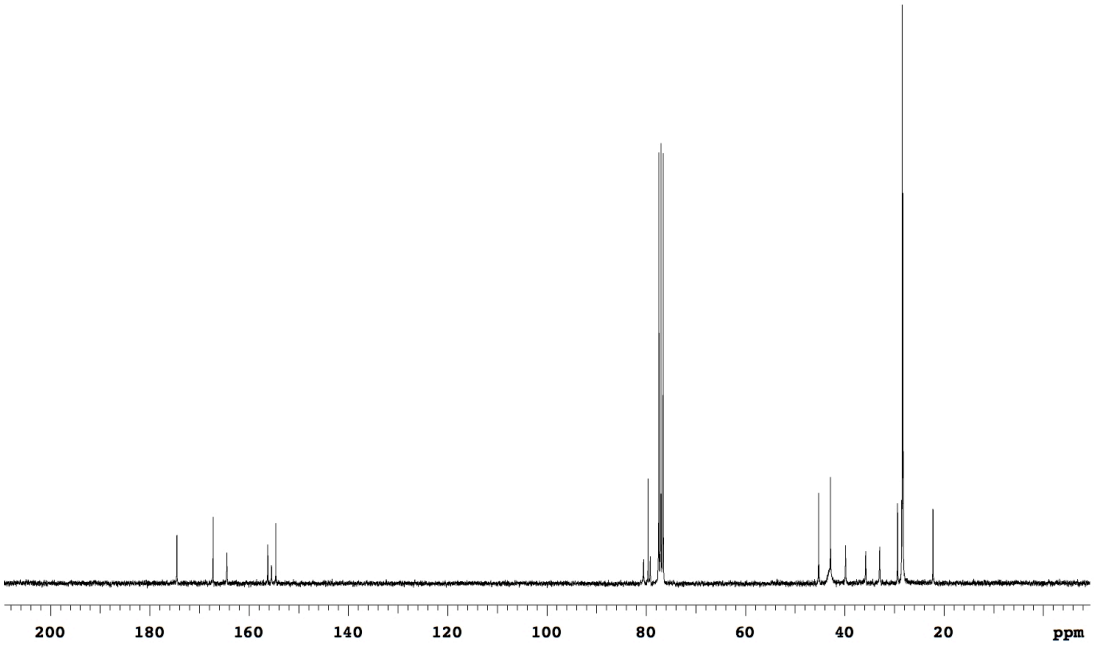
Compound **12f** was prepared from (5.41 g, 10.1 mmol) of **11f**. Flash chromatography (30 % to 50 % EtOAc/Hexanes) afforded 5.10 g (83 %) **12f** as a white solid.

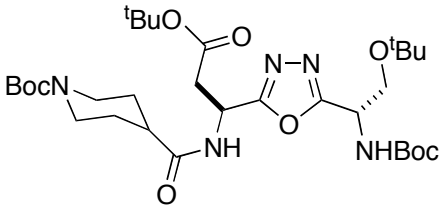
^{1}H NMR (300MHz, $CDCl_3$) δ 6.80 (d, 1H, J = 7.8 Hz), 5.44 (br, 1H), 5.35-5.21 (m, 1H), 4.71 (br, 1H), 4.49 (d, 2H, J = 5.9 Hz), 4.07 (br, 2H), 3.14-2.96 (m, 2H), 2.80-2.60 (m, 2H), 2.40-2.24 (m, 1H), 2.01-1.56 (m, 6H), 1.54-1.25 (m, 31H)

^{13}C NMR (75MHz, $CDCl_3$) δ 174.5, 167.2, 164.5, 156.2, 155.5, 154.6, 80.5, 79.6, 79.2, 77.2, 45.2, 42.9, 39.8, 35.8, 33.0, 29.3, 28.5, 28.4, 28.4, 28.2, 22.2

MS (ESI, m/z) 633.2 $(M+Na)^+$







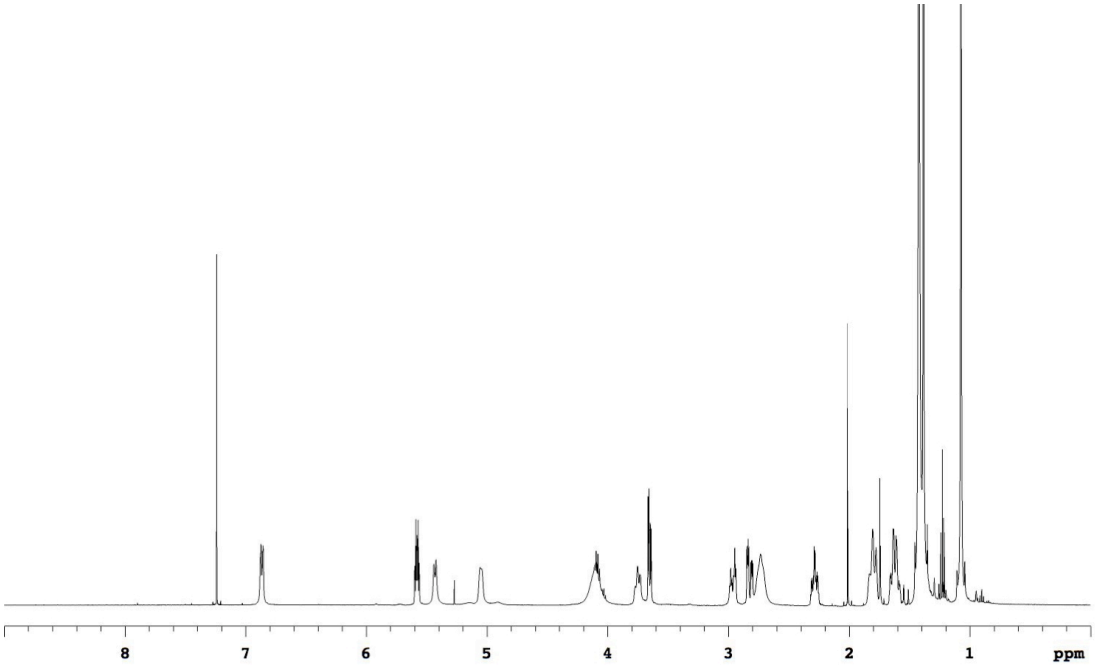
12g

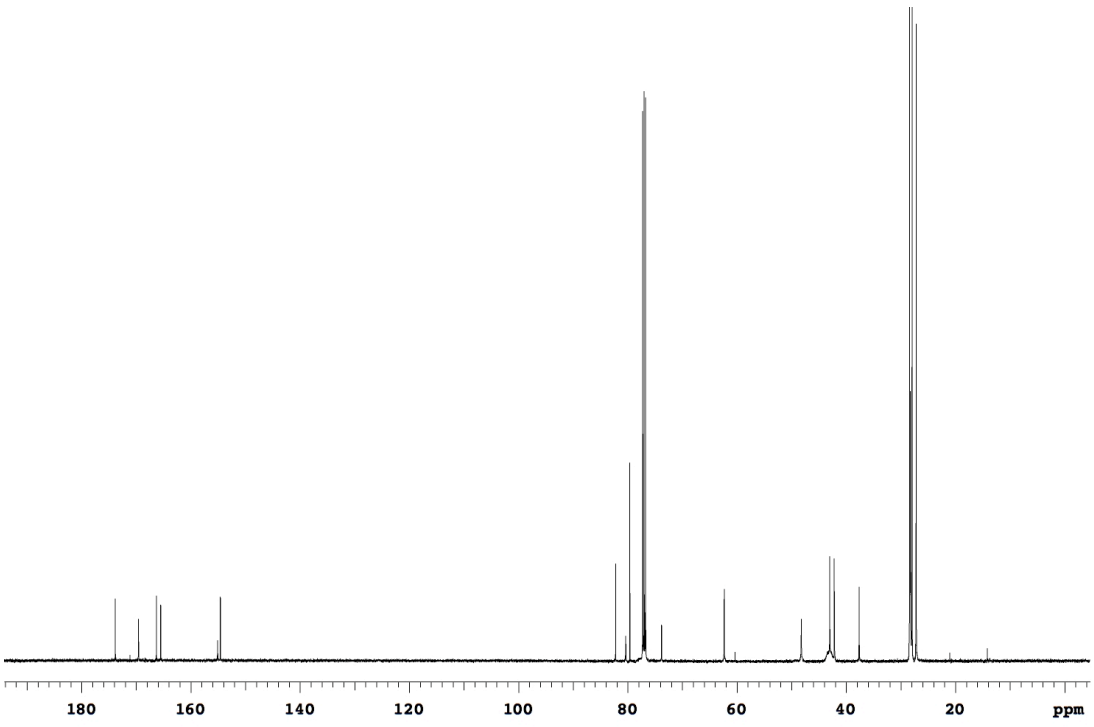
Compound **12g** was prepared from (3.62 g, 6.4 mmol) of **11g**. Flash chromatography (30 % to 50 % EtOAc/Hexanes) afforded 3.50 g (84 %) **12g** as a white solid.

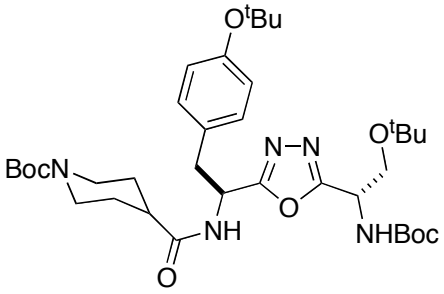
^{1}H NMR (500 MHz, $CDCl_3$) δ 6.91-6.82 (m, 1H), 5.58 (dt, 1H, J = 4.9, 8.8 Hz), 5.43 (d, 1H, J = 8.8 Hz), 5.06 (br, 1H), 4.10 (br, 2H), 3.80-3.71 (m, 1H), 3.65 (dd, 1H, J = 4.1, 9.2 Hz), 2.97 (dt, 1H, J = 4.5, 16.5 Hz), 2.82 (ddd, 1H, J = 2.6, 5.2, 16.5 Hz), 2.73 (br, 2H), 2.33-2.25 (m, 1H), 1.87-1.76 (m, 2H), 1.68-1.57 (m, 2H), 1.42 (br, 18H), 1.39 (s, 9H), 1.08 (s, 9H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 173.8, 169.6, 166.3, 165.5, 155.1, 154.6, 82.2, 80.4, 79.6, 73.8, 62.3, 48.2, 43.0, 42.2, 37.6, 28.4, 28.4, 28.3, 28.2, 27.9, 27.2

MS (MALDI, m/z) 661.9 $(M+Na)^+$


8
7
6
5
4
3
2
1
ppm


180
160
140
120
100
80
60
40
20
ppm



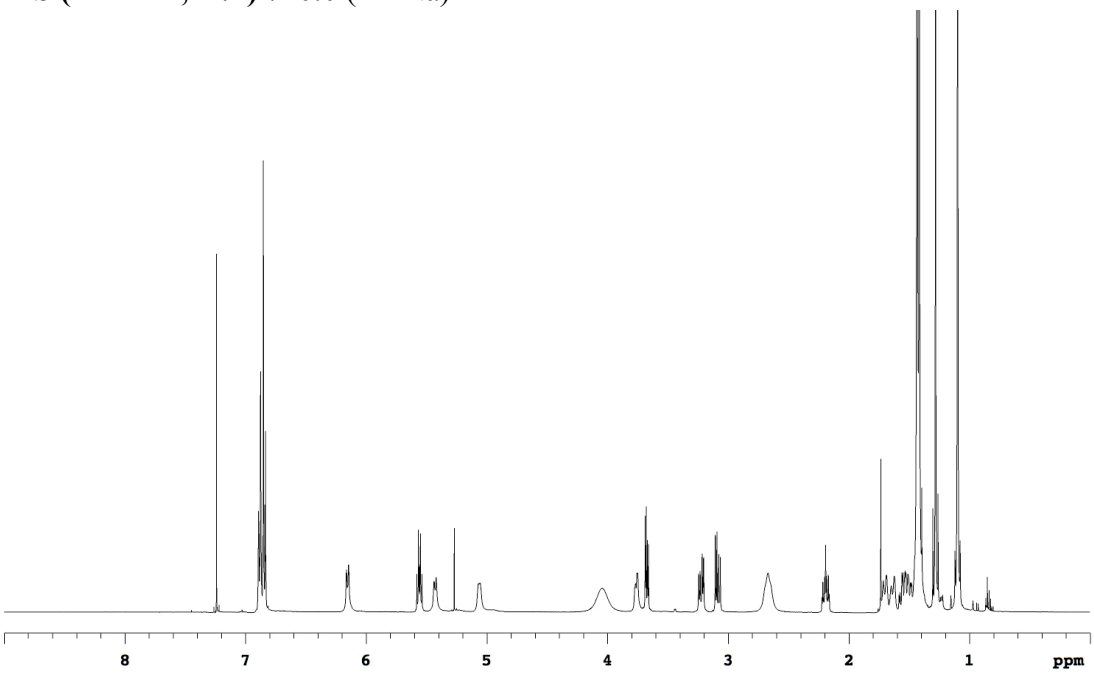
12h

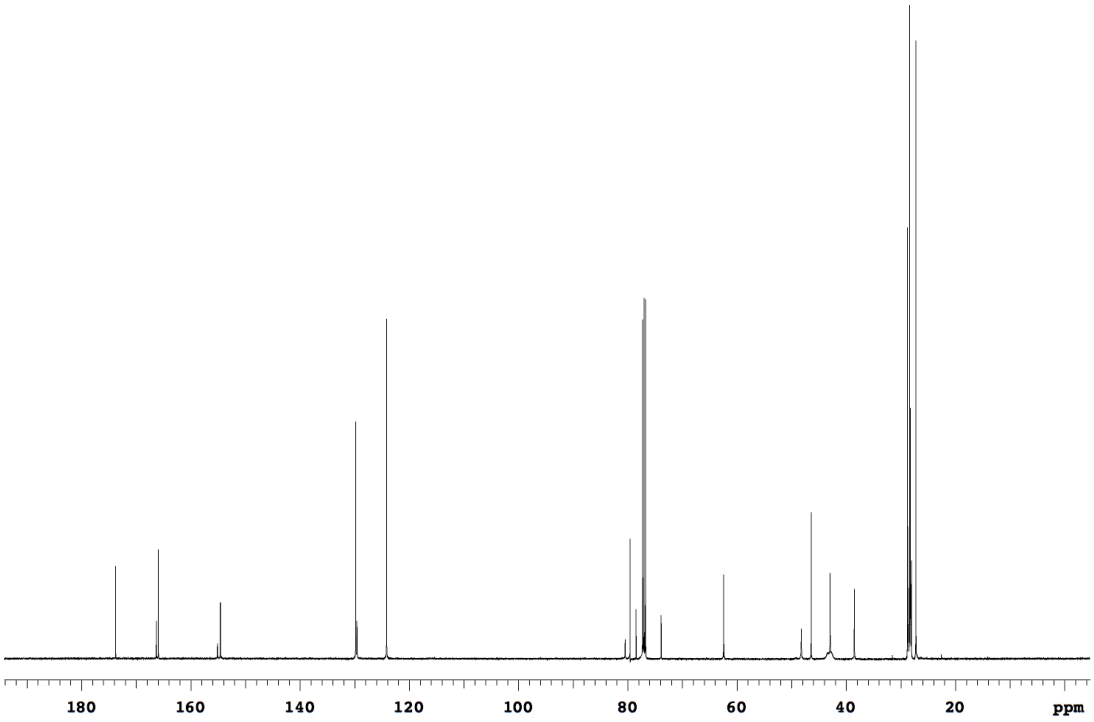
Compound **12h** was prepared from (4.12 g, 6.7 mmol) of **11h**. Flash chromatography (30 % to 50 % EtOAc/Hexanes) afforded 3.70 g (80 %) **12h** as a white solid.

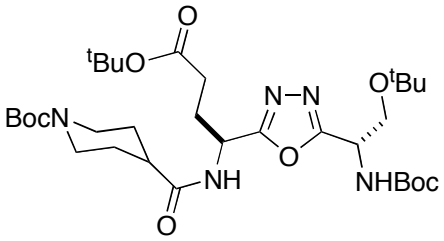
^{1}H NMR (500 MHz, $CDCl_3$) δ 6.92-6.81 (m, 4H), 6.15 (d, 1H, J = 8.4 Hz), 5.59-5.53(m, 1H), 5.43 (d, 1H, J = 8.4 Hz), 5.11-5.02 (m, 1H), 4.04 (br, 2H), 3.80-3.72 (m, 1H), 3.23 (dd, 1H, J = 6.0, 14.1 Hz), 3.09 (dd, 1H, J = 6.5, 14.1 Hz), 2.67 (dd, 1H, J = 4.1, 9.3 Hz), 2.19 (tt, 1H, J = 3.7, 11.5 Hz), 1.75-1.47 (m, 4H), 1.43 (s, 9H), 1.42 (s, 9H), 1.28 (s, 9H), 1.10 (s, 9H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 173.8, 166.3, 165.9, 155.1, 154.6, 154.5, 129.8, 129.6, 124.2, 80.5, 79.6, 78.5, 73.9, 62.4, 48.2, 46.4, 42.9, 38.5, 28.8, 28.6, 28.4, 28.3, 28.1, 27.2

MS (MALDI, m/z) 710.0 $(M+Na)^+$







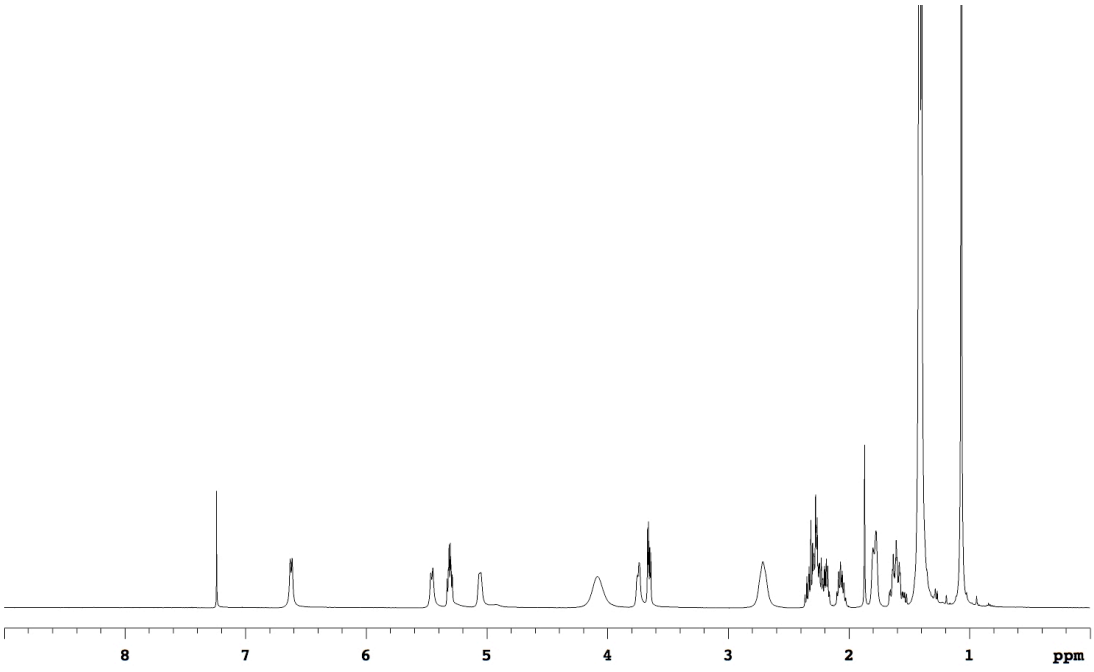
12i

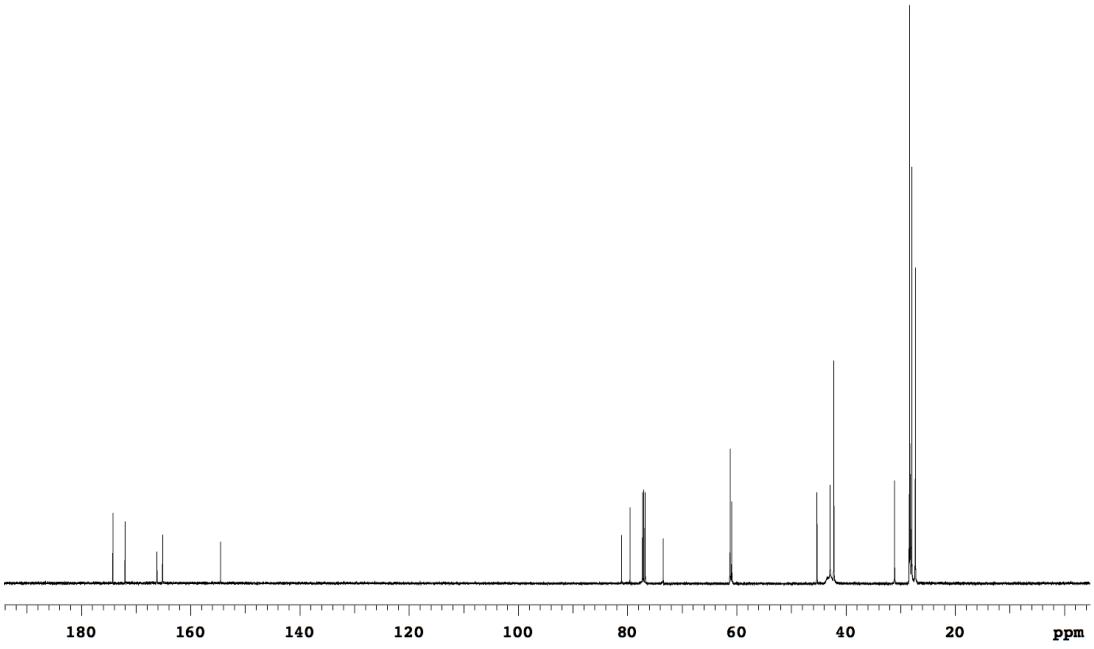
Compound **12i** was prepared from (7.02 g, 12.2 mmol) of **11i**. Flash chromatography (30 % to 50 % EtOAc/Hexanes) afforded 8.30 g (82 %) **12i** as a white solid.

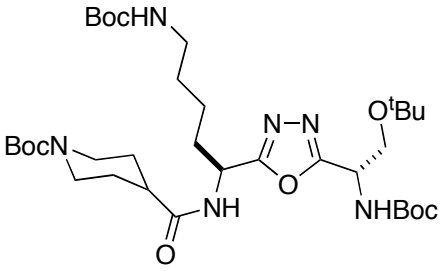
^{1}H NMR (500 MHz, $CDCl_3$) δ 6.62 (d, 1H, J = 8.0 Hz), 5.46 (d, 1H, J = 8.5 Hz), 5.35-5.26(m, 1H), 5.06 (br, 1H), 4.09 (br, 2H), 3.78-3.70 (br, 1H), 3.65 (dd, 1H, J = 4.0, 9.2 Hz), 2.71 (br, 2H), 2.38-2.15 (m, 4H), 2.11-2.01 (m, 1H), 1.83-1.74 (m, 2H), 1.68-1.54 (m, 2H), 1.49-1.31 (m, 27H), 1.07 (s, 9H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 174.5, 172.3, 166.5, 155.4, 154.9, 81.4, 80.7, 79.8, 74.1, 62.8, 48.5, 45.4, 43.2, 31.4, 28.7, 28.7, 28.6, 28.5, 28.3, 27.5

MS (MALDI, m/z) 676.1 $(M+Na)^+$


8
7
6
5
4
3
2
1
ppm


180
160
140
120
100
80
60
40
20
ppm



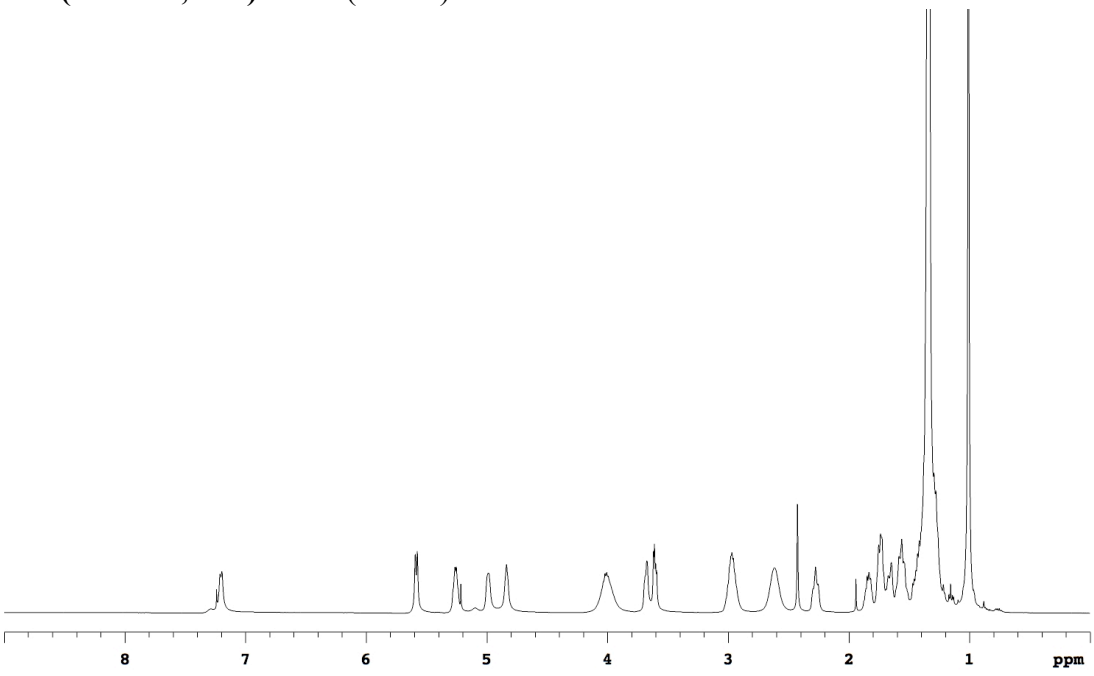
12j

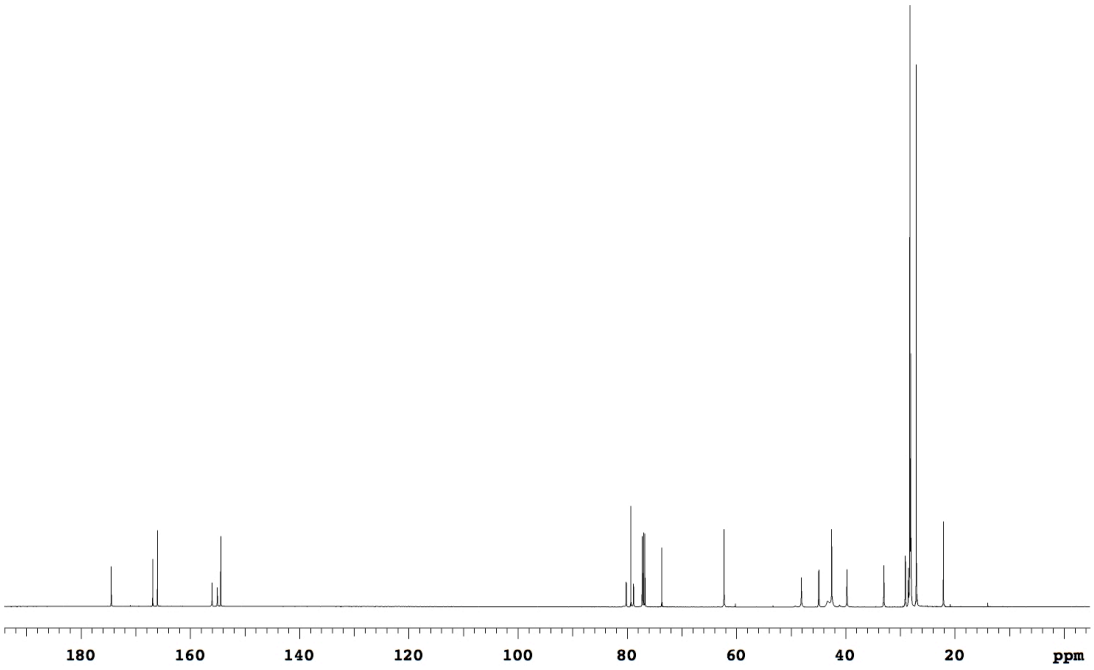
Compound **12j** was prepared from (6.54 g, 10.6 mmol) of **11j**. Flash chromatography (30 % to 50 % EtOAc/Hexanes) afforded 5.60 g (77 %) **12j** as a white solid.

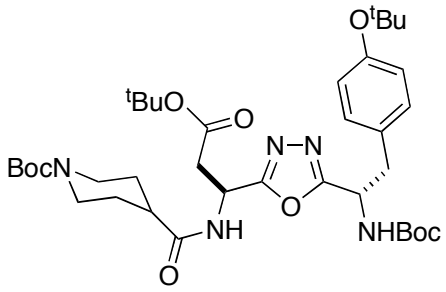
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.20 (d, 1H, J = 7.6 Hz), 5.59 (d, 1H, J = 8.5 Hz), 5.32-5.19(m, 1H), 4.99 (br, 1H), 4.84 (br, 1H), 4.01 (br, 2H), 3.73-3.55 (m, 2H), 3.05-2.87 (m, 2H), 2.62 (br, 2H), 2.33-2.21 (m, 1H), 1.90-1.50 (m, 6H), 1.49-1.23 (m, 31H), 1.01 (s, 9H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 174.5, 166.9, 166.0, 156.0, 155.1, 154.4, 80.2, 79.3, 78.8, 73.7, 62.3, 48.1, 44.9, 42.6, 39.8, 33.0, 29.1, 28.5, 28.3, 28.2, 28.1, 28.0, 27.1, 22.1

MS (MALDI, m/z) 719.1 $(M+Na)^+$







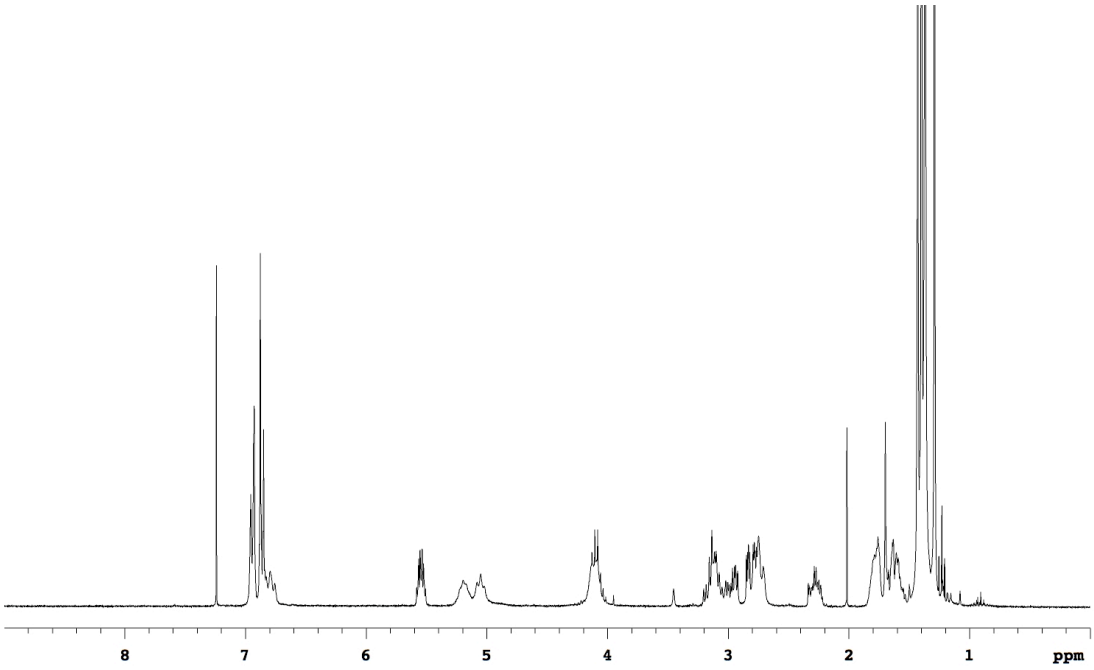
12k

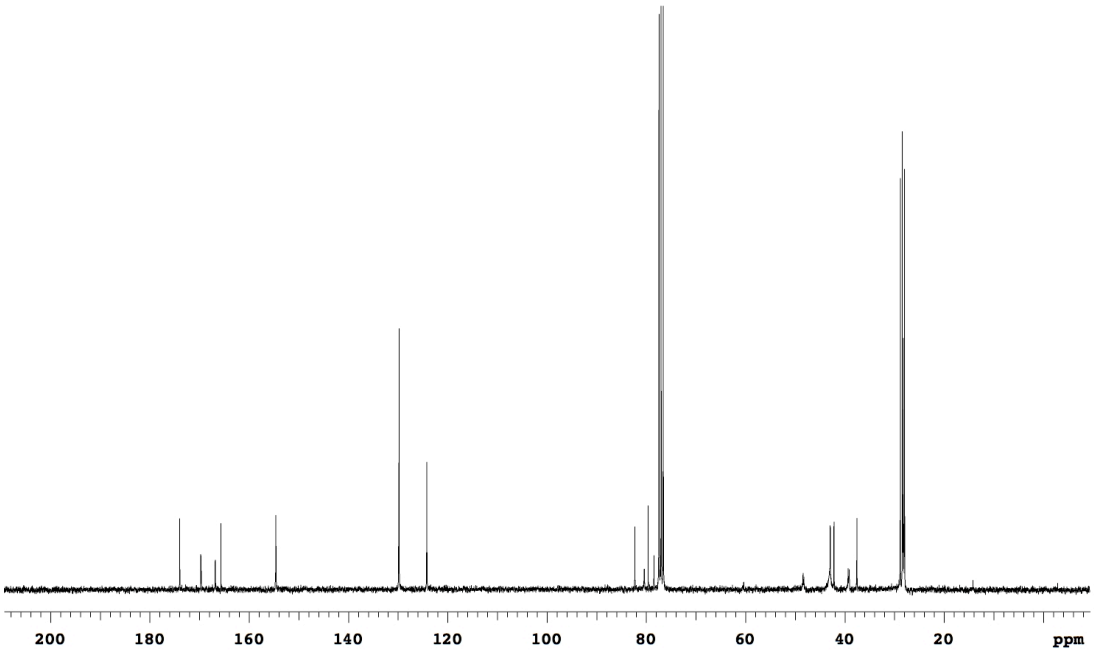
Compound **12k** was prepared from (3.35 g, 5.2 mmol) of **11k**. Flash chromatography (30 % to 50 % EtOAc/Hexanes) afforded 2.80 g (74 %) **12k** as a white solid.

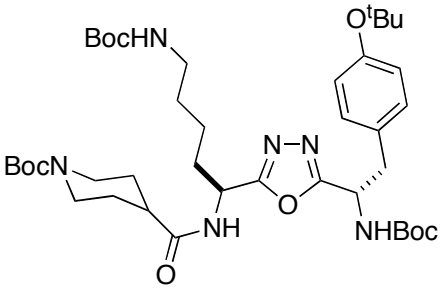
^{1}H NMR (300MHz, $CDCl_3$) δ 6.99-6.72 (m, 4H), 5.59-5.50 (m, 2H), 5.28-4.96(m, 2H), 4.11 (br, 2H), 3.23-2.90 (m, 3H), 2.87-2.88 (m, 3H), 2.35-2.21 (m, 1H), 1.86-1.52 (m, 4H), 1.48-1.32 (m, 27H), 1.29 (s, 9H)

^{13}C NMR (75MHz, $CDCl_3$) δ 174.0, 169.8, 169.7, 166.9, 166.8, 165.7, 154.6, 129.9, 129.8, 124.2, 82.3, 80.4, 79.6, 78.5, 48.4, 43.0, 42.2, 39.3, 37.6, 28.8, 28.4, 28.4, 28.3, 28.2, 28.0

MS (MALDI, m/z) 738.1 $(M+Na)^+$


8
7
6
5
4
3
2
1
ppm


200
180
160
140
120
100
80
60
40
20
ppm



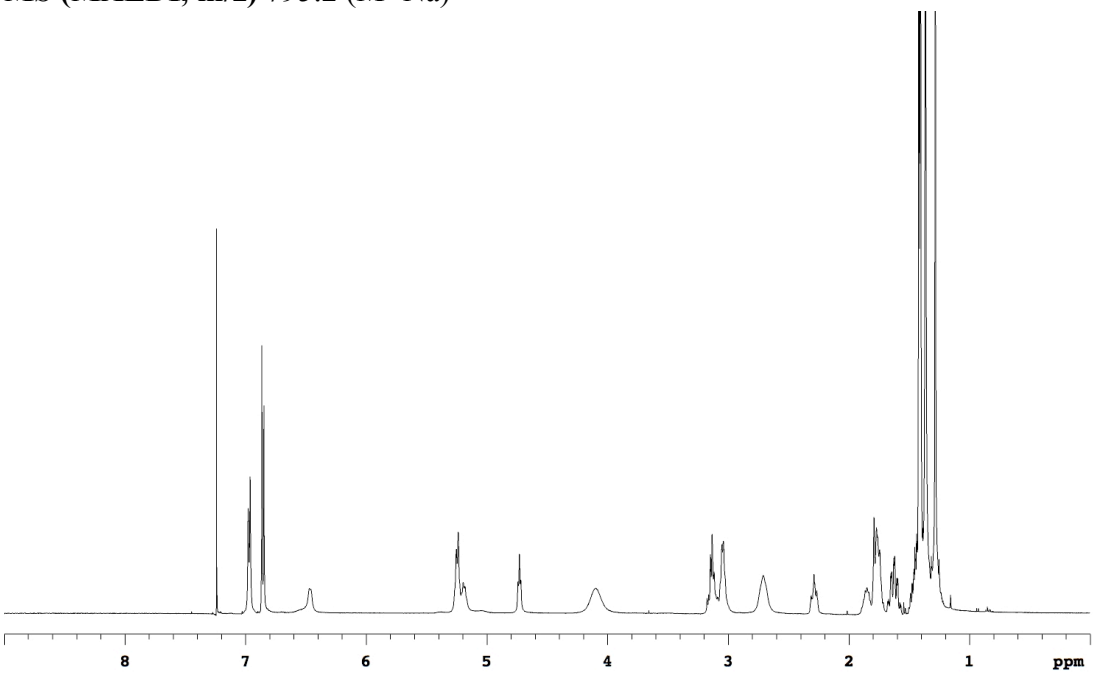
12l

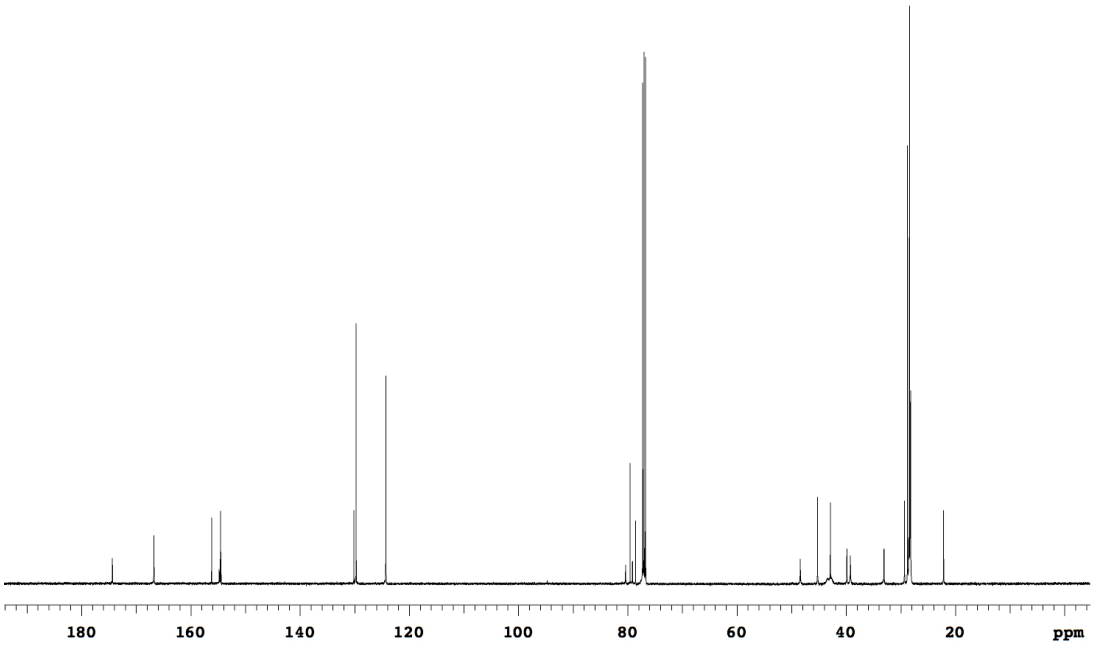
Compound **12l** was prepared from (5.96 g, 8.6 mmol) of **11l**. Flash chromatography (30 % to 50 % EtOAc/Hexanes) afforded 5.20 g (78 %) **12l** as a white solid.

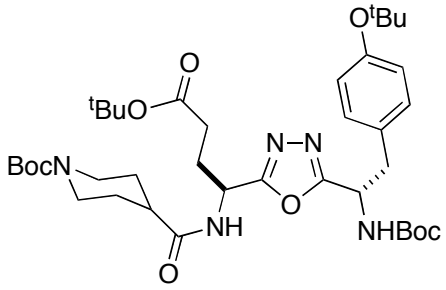
^{1}H NMR (500 MHz, $CDCl_3$) δ 6.97 (d, 2H, J = 8.3 Hz), 6.86 (d, 2H, J = 8.3 Hz), 6.54-6.41(br, 1H), 5.30-5.13 (m, 3H), 4.73 (t, 1H, J = 5.7 Hz), 4.10 (br, 2H), 3.20-2.98 (m, 4H), 2.71 (br, 2H), 2.34-2.24 (m, 1H), 1.92-1.69 (m, 4H), 1.51-1.20 (m, 40H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 174.4, 166.8, 156.2, 154.8, 154.6, 154.5, 130.1, 129.7, 124.3, 80.4, 79.6, 79.1, 78.5, 48.4, 45.2, 42.9, 39.9, 39.2, 33.1, 29.3, 28.8, 28.6, 28.4, 28.4, 28.3, 28.2, 22.2

MS (MALDI, m/z) 795.2 $(M+Na)^+$







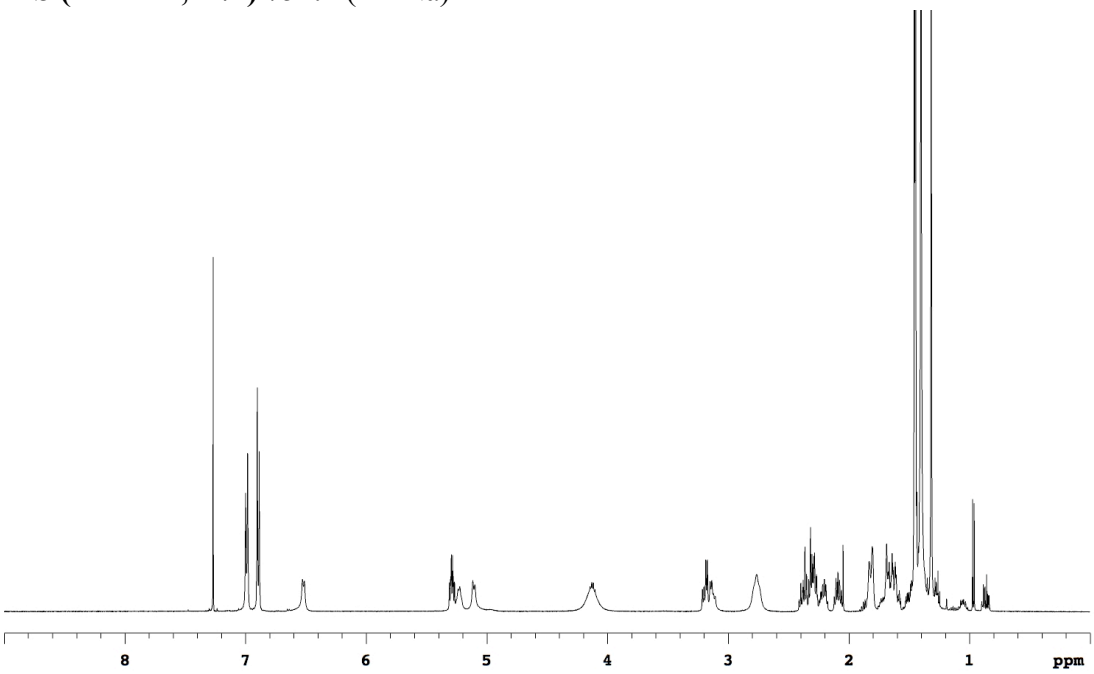
12m

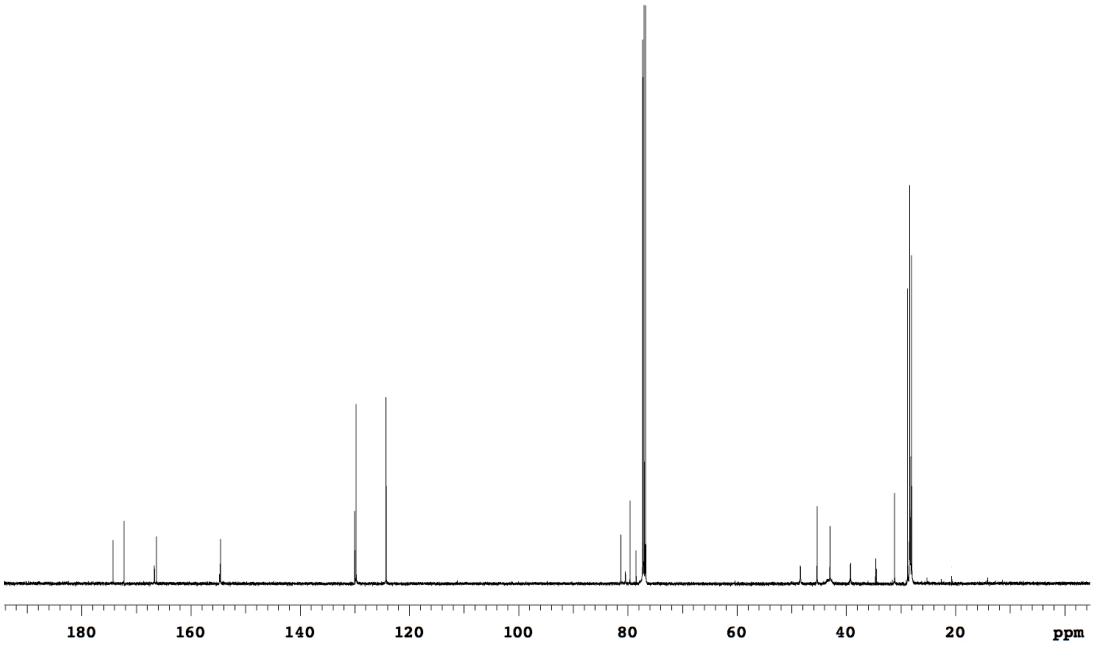
Compound **12m** was prepared from (4.58 g, 7.0 mmol) of **11m**. Flash chromatography (30 % to 50 % EtOAc/Hexanes) afforded 4.60 g (89 %) **12m** as a white solid.

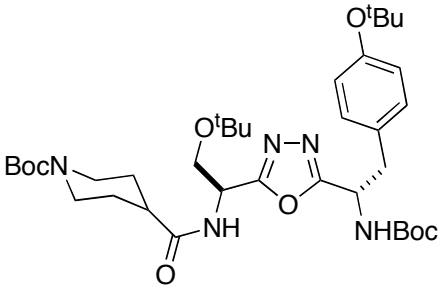
^{1}H NMR (500 MHz, $CDCl_3$) δ 6.99 (d, 2H, J = 8.3 Hz), 6.89 (d, 2H, J = 8.3 Hz), 6.52 (d, 1H, J = 7.6 Hz), 5.32-5.18 (m, 2H), 5.11 (d, 1H, J = 8.4 Hz), 4.13 (br, 2H), 3.19 (dd, 1H, J = 6.5, 14.0 Hz), 3.13 (dd, 1H, J = 6.6, 14.0 Hz), 2.77 (br, 2H), 2.43-2.16 (m, 4H), 2.14-2.04 (m, 1H), 1.90-1.56 (m, 4H), 1.50-1.34 (m, 27H), 1.32 (s, 9H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 174.3, 172.2, 166.7, 166.3, 154.7, 154.6, 154.6, 130.0, 129.8, 124.3, 81.3, 80.4, 79.6, 78.5, 48.4, 45.3, 43.0, 39.2, 34.6, 31.2, 28.8, 28.5, 28.4, 28.3, 28.2, 28.0

MS (MALDI, m/z) 752.1 $(M+Na)^+$







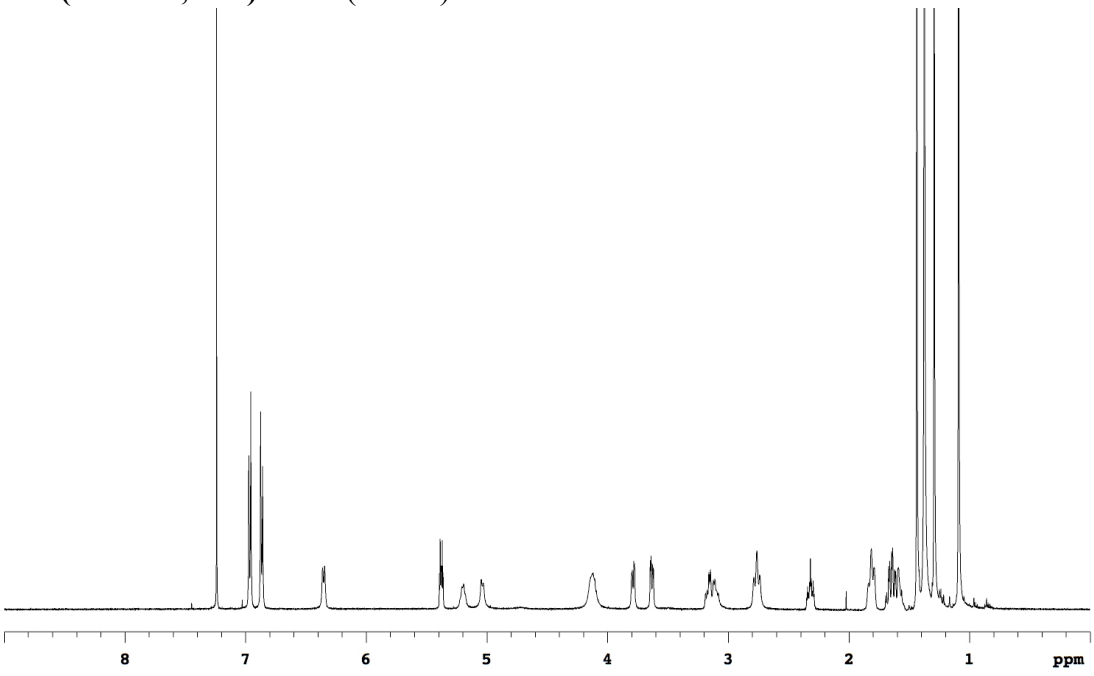
12n

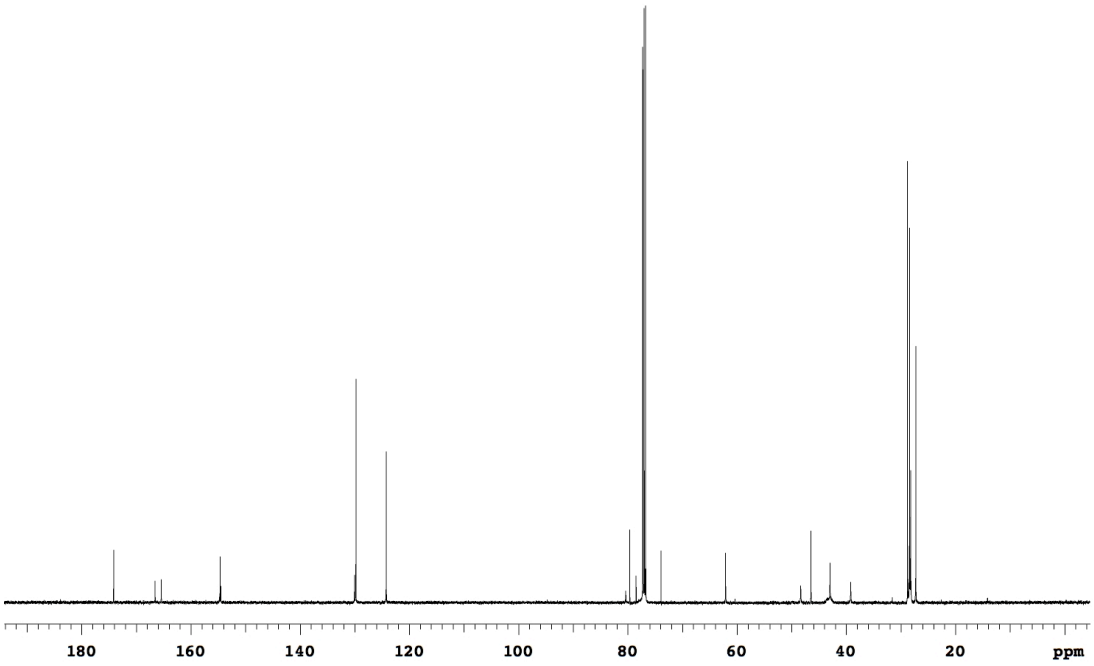
Compound **12n** was prepared from (3.63 g, 5.9 mmol) of **11n**. Flash chromatography (30 % to 50 % EtOAc/Hexanes) afforded 3.50 g (85 %) **12n** as a white solid.

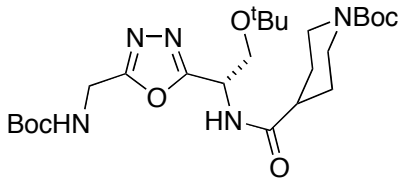
^{1}H NMR (500 MHz, $CDCl_3$) δ 6.96 (d, 2H, J = 8.4 Hz), 6.87 (d, 2H, J = 8.4 Hz), 6.35 (d, 1H, J = 8.3 Hz), 5.38 (dt, 1H, J = 3.5, 8.4 Hz), 5.25-5.15 (m, 1H), 5.04 (d, 1H, J = 8.3 Hz), 4.12 (br, 2H), 3.79 (dd, 1H, J = 2.7, 9.2 Hz), 3.63 (dd, 1H, J = 3.6, 9.2 Hz), 3.21-3.04 (m, 2H), 2.76 (t, 2H, J = 12.2 Hz), 2.32 (tt, 1H, J = 3.7, 11.5 Hz), 1.86-1.76 (m, 2H), 1.71-1.54 (m, 2H), 1.44 (s, 9H), 1.38 (s, 9H), 1.29 (s, 9H), 1.09 (s, 9H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 174.2, 166.6, 165.4, 154.7, 154.6, 154.5, 130.0, 129.8, 124.2, 80.4, 79.6, 78.5, 73.9, 62.1, 48.3, 46.5, 43.0, 39.2, 28.8, 28.5, 28.4, 28.3, 28.2, 27.3

MS (MALDI, m/z) 710.1 $(M+Na)^+$







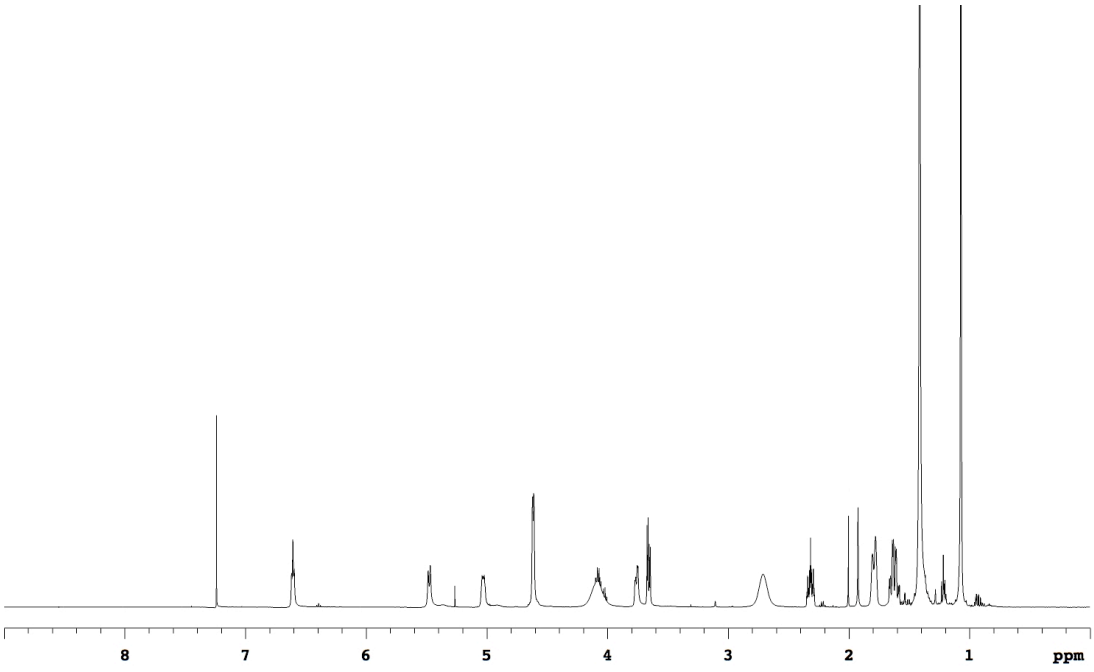
12o

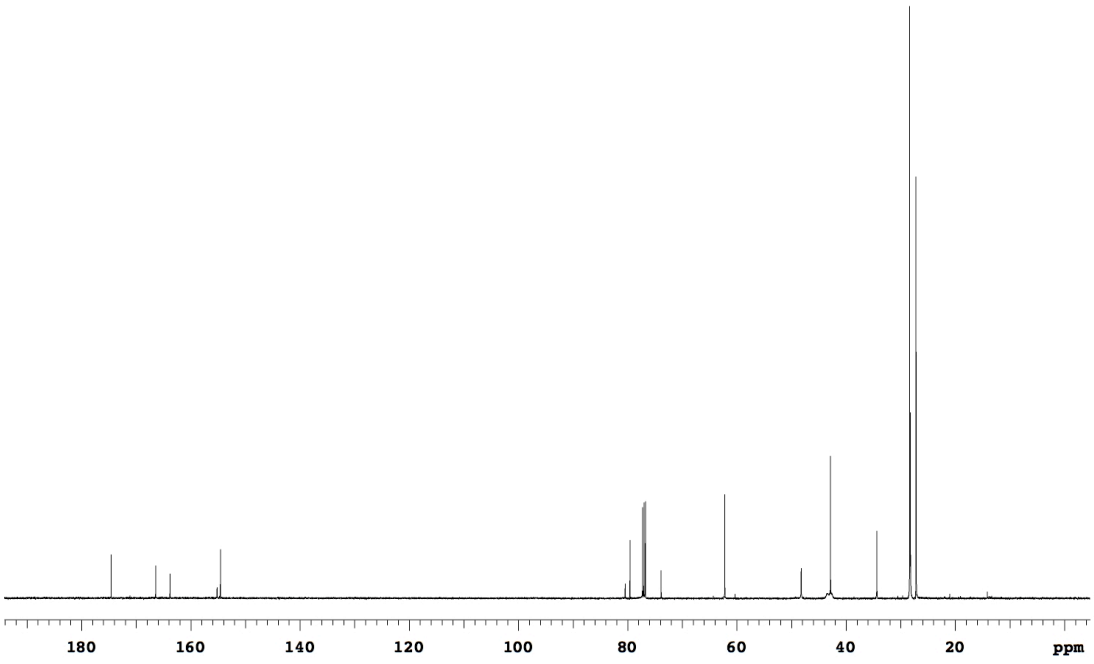
Compound **12o** was prepared from (0.36 g, 0.8 mmol) of **11o**. Flash chromatography (30 % to 70 % EtOAc/Hexanes) afforded 0.31 g (74 %) **12o** as a white solid.

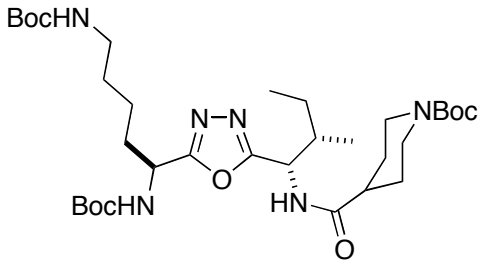
^{1}H NMR (500 MHz, $CDCl_3$) δ 6.61 (t, 1H, J = 5.4 Hz), 5.48 (d, 1H, J = 8.6 Hz), 5.07-4.99 (m, 1H), 4.62 (d, 2H, J = 5.4 Hz), 3.76 (dd, 1H, J = 3.2, 9.3 Hz), 3.66 (dd, 1H, J = 4.0, 9.3 Hz), 2.71 (br, 2H), 2.32 (tt, 1H, J = 3.7, 11.5 Hz), 1.84-1.74 (m, 2H), 1.69-1.57 (m, 2H), 1.41 (s, 18 H), 1.07 (s, 9H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 174.6, 166.4, 163.8, 155.2, 154.6, 80.4, 79.6, 73.9, 62.2, 48.2, 42.8, 34.4, 28.4, 28.2, 27.2

MS (MALDI, m/z) 547.8 $(M+Na)^+$


8
7
6
5
4
3
2
1
ppm


180
160
140
120
100
80
60
40
20
ppm



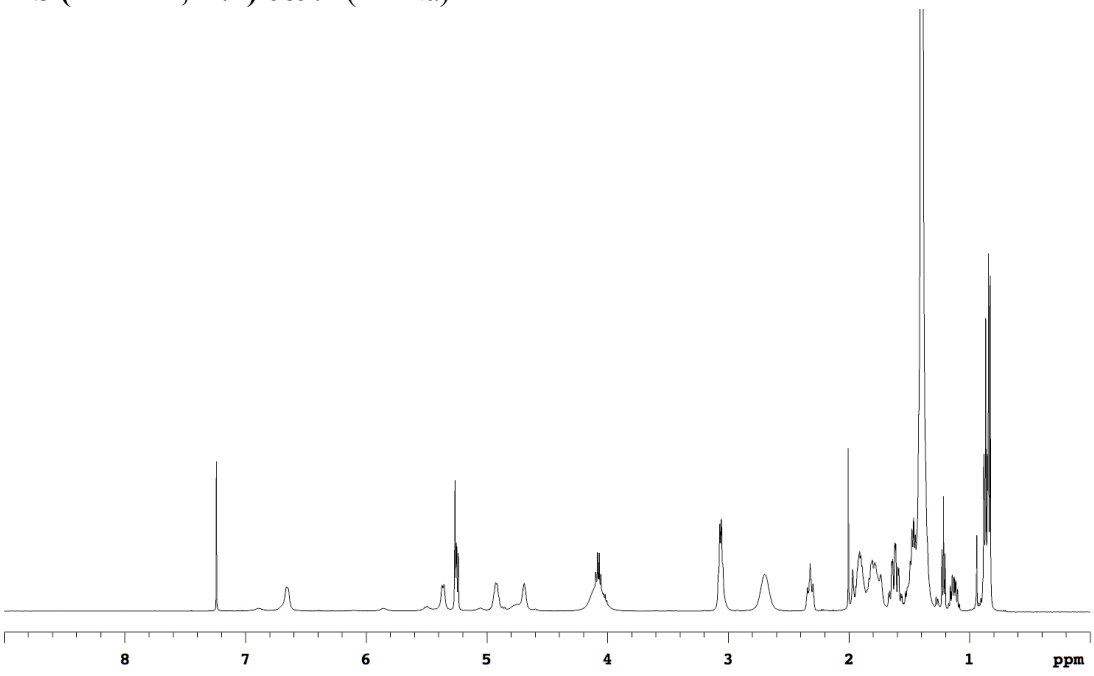
12p

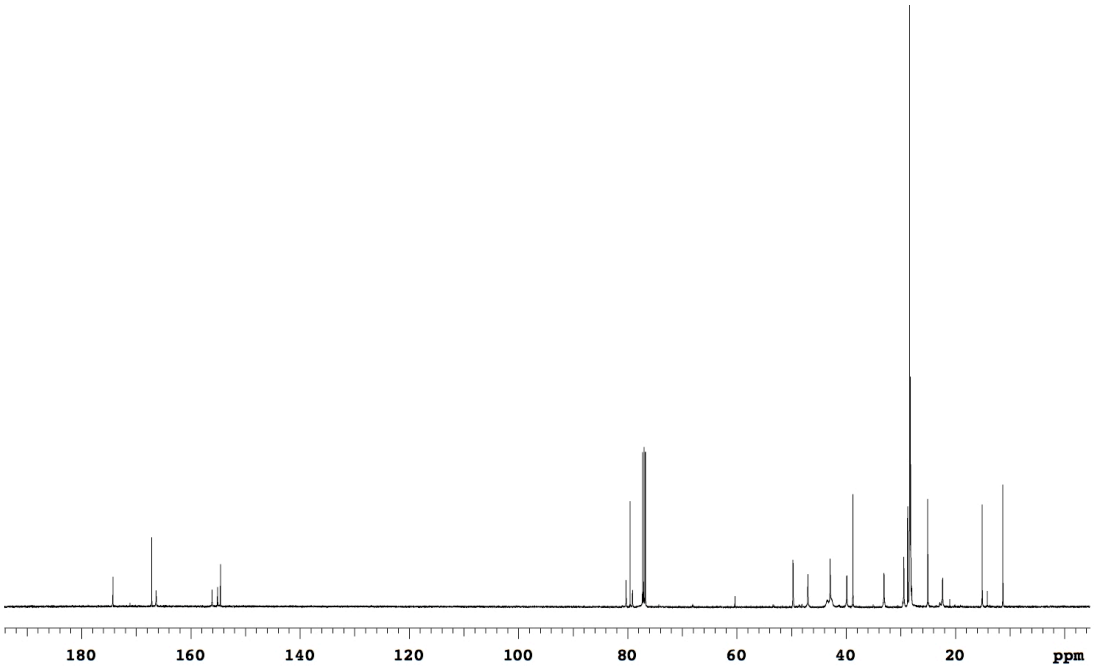
Compound **12p** was prepared from (1.18 g, 1.9 mmol) of **11p**. Flash chromatography (30 % to 70 % EtOAc/Hexanes) afforded 1.16 g (92 %) **12p** as a white solid.

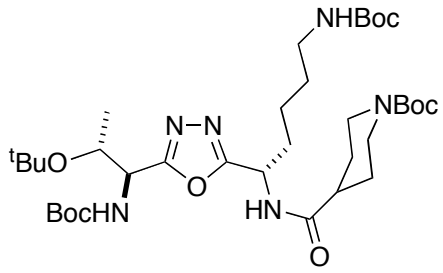
^{1}H NMR (500 MHz, $CDCl_3$) δ 6.65 (d, 1H, J = 7.2 Hz), 5.36 (d, 1H, J = 7.8 Hz), 5.30-5.21 (m, 1H), 4.93 (br, 1H), 4.69 (br, 1H), 4.08 (br, 2H), 3.13-2.99 (m, 2H), 2.70 (br, 2H), 2.32 (tt, 1H, J = 3.7, 11.5 Hz), 1.99-1.70 (m, 5H), 1.69-1.55 (m, 2H), 1.54-1.29 (m, 32 H), 1.19-1.07 (m, 1H), 0.91-0.80 (m, 6H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 174.3, 167.2, 166.3, 156.1, 155.1, 154.6, 80.3, 79.6, 79.1, 49.7, 47.0, 42.9, 39.9, 38.7, 33.1, 29.4, 28.7, 28.4, 28.3, 28.2, 28.1, 25.0, 22.3, 15.1, 11.3

MS (MALDI, m/z) 689.1 $(M+Na)^+$







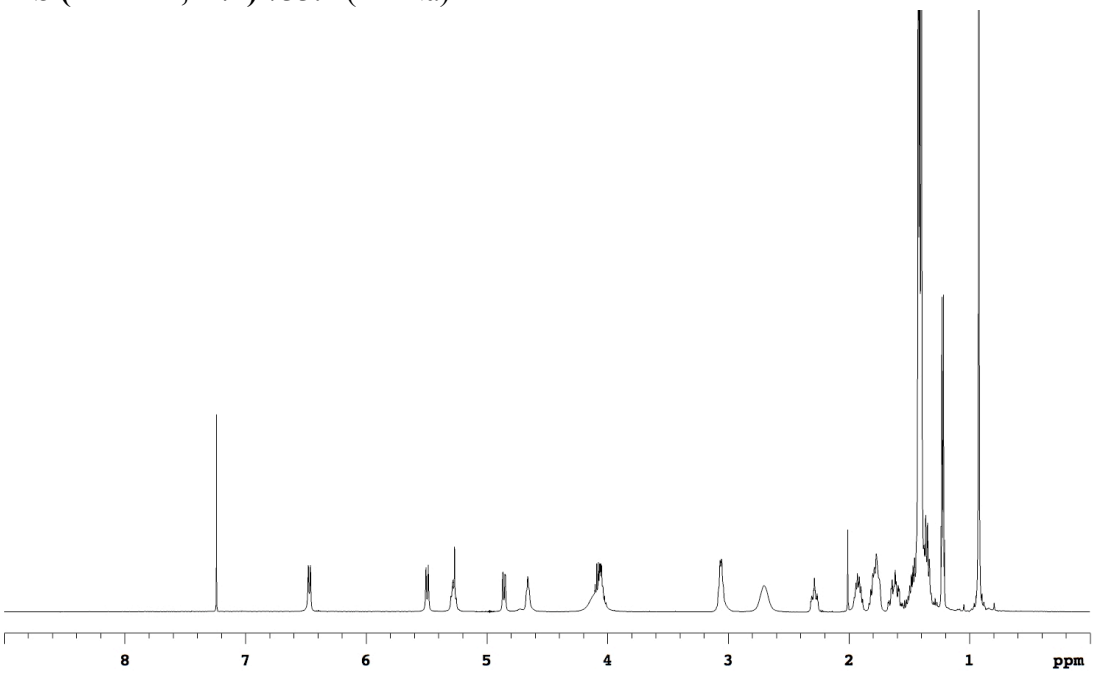
12q

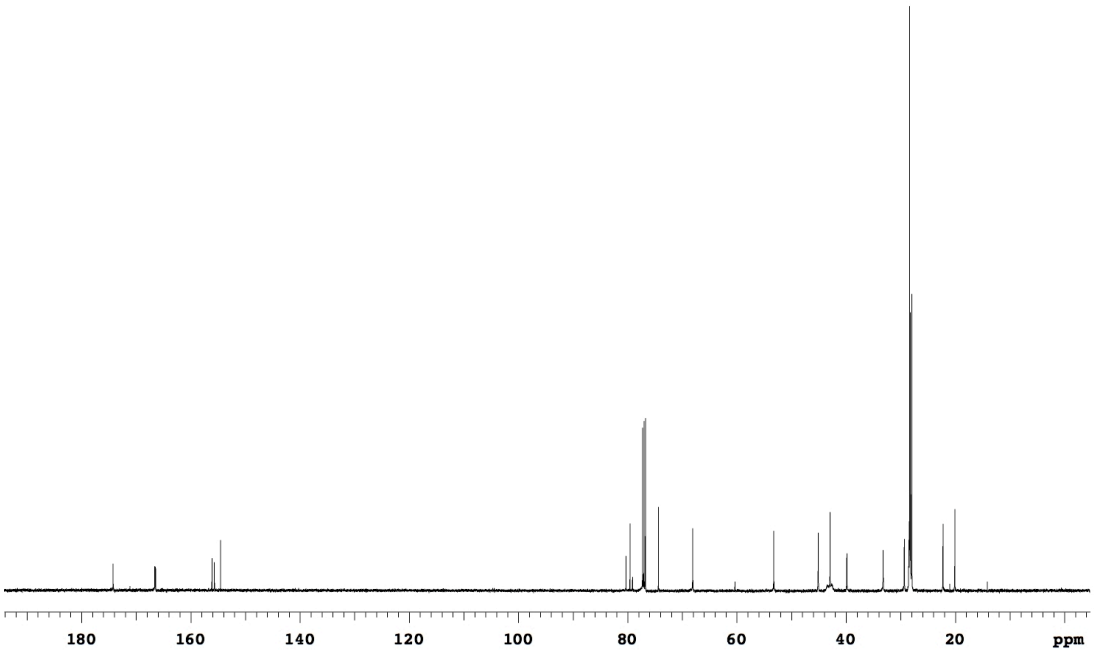
Compound **12q** was prepared from (1.70 g, 2.7 mmol) of **11q**. Flash chromatography (30 % to 70 % EtOAc/Hexanes) afforded 1.35 g (70 %) **12q** as a white solid.

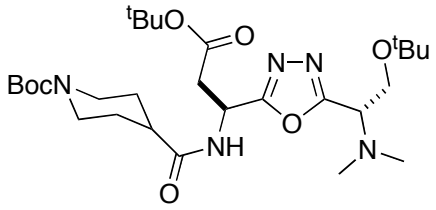
^{1}H NMR (500 MHz, $CDCl_3$) δ 6.47 (d, 1H, J = 8.5 Hz), 5.49 (d, 1H, J = 9.5 Hz), 5.33-5.22 (m, 1H), 4.86 (dd, 1H, J = 1.7, 9.5 Hz), 4.70-4.60 (m, 1H), 4.22-3.96 (m, 3H), 3.13-2.98 (m, 2H), 2.70 (br, 2H), 2.29 (tt, 1H, J = 3.7, 11.5 Hz), 1.98-1.86 (m, 2H), 1.85-1.71 (m, 2H), 1.69-1.55 (m, 2 H), 1.54-1.21 (m, 31H), 1.22 (d, 3H, J = 6.2 Hz), 0.92 (s, 9H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 174.2, 166.6, 166.5, 156.1, 155.7, 154.6, 80.3, 79.6, 79.1, 74.4, 68.1, 53.3, 45.1, 42.9, 39.9, 33.2, 29.3, 28.7, 28.5, 28.4, 28.4, 28.3, 28.0, 22.3, 20.1

MS (MALDI, m/z) 733.1 $(M+Na)^+$







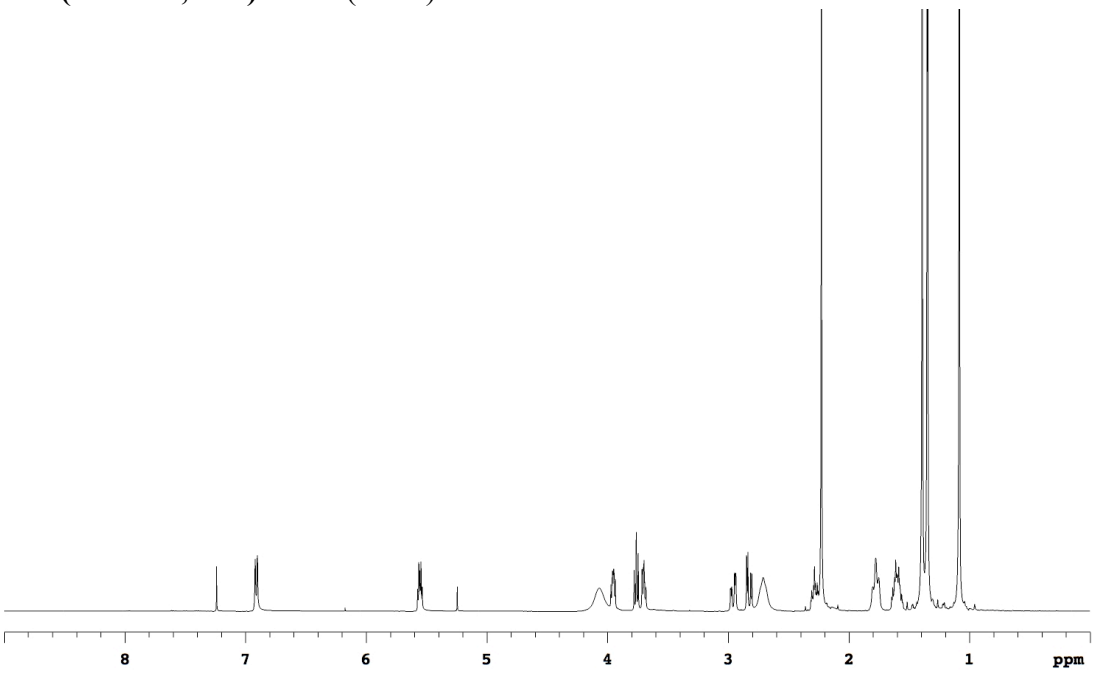
12r

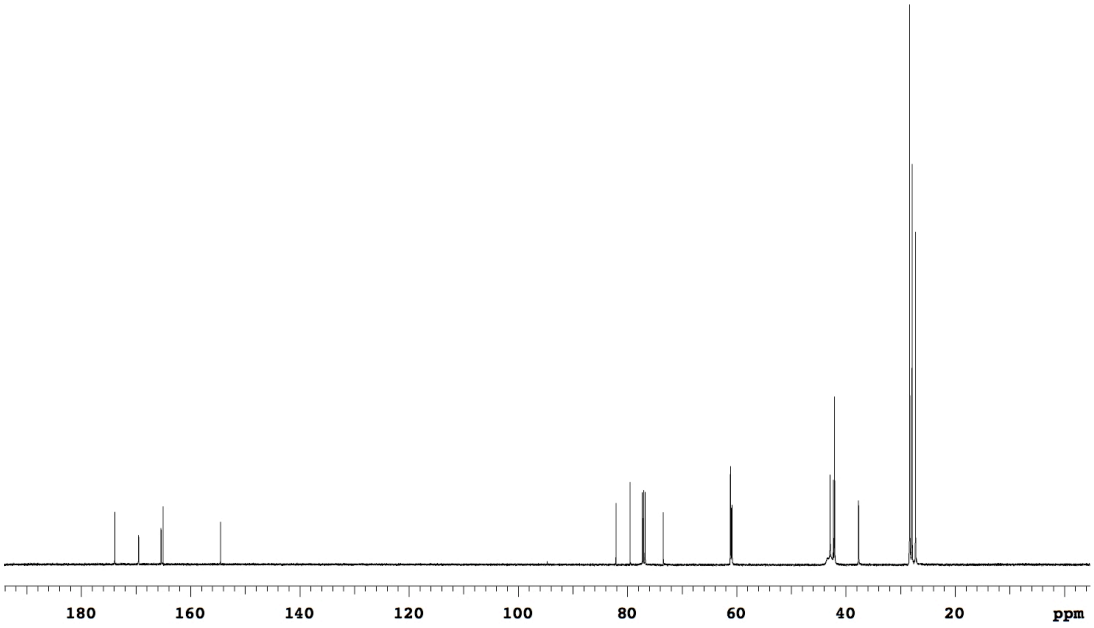
Compound **12r** was prepared from (0.93 g, 2.6 mmol) of **11r**. Flash chromatography (5% MeOH/CH_2Cl_2) afforded 1.21 g (82 %) **12r** as a white solid.

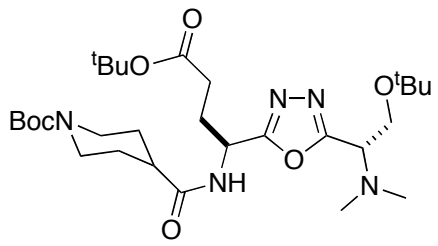
^{1}H NMR (500 MHz, $CDCl_3$) δ 6.91 (d, 1H, J = 8.8 Hz), 5.55 (dt, 1H, J = 5.0, 8.8 Hz), 4.16-3.92 (m, 3H), 3.80-3.66 (m, 2H), 2.96 (ddd, 1H, J = 1.5, 4.7, 16.6 Hz), 2.83 (dd, 1H, J = 5.1, 16.6 Hz), 2.71 (br, 2H), 2.28 (tt, 1H, J = 3.7, 11.5 Hz), 2.23 (s, 6H), 1.84-1.72 (m, 2H), 1.66-1.54 (m, 2H), 1.45-1.28 (m, 18H), 1.09 (s, 9H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 173.9, 169.6, 165.4, 165.0, 154.5, 82.1, 79.5, 73.4, 61.2, 61.0, 60.8, 42.9, 42.3, 42.3, 42.1, 37.7, 28.3, 27.9, 27.2

MS (MALDI, m/z) 568.0 $(M+H)^+$







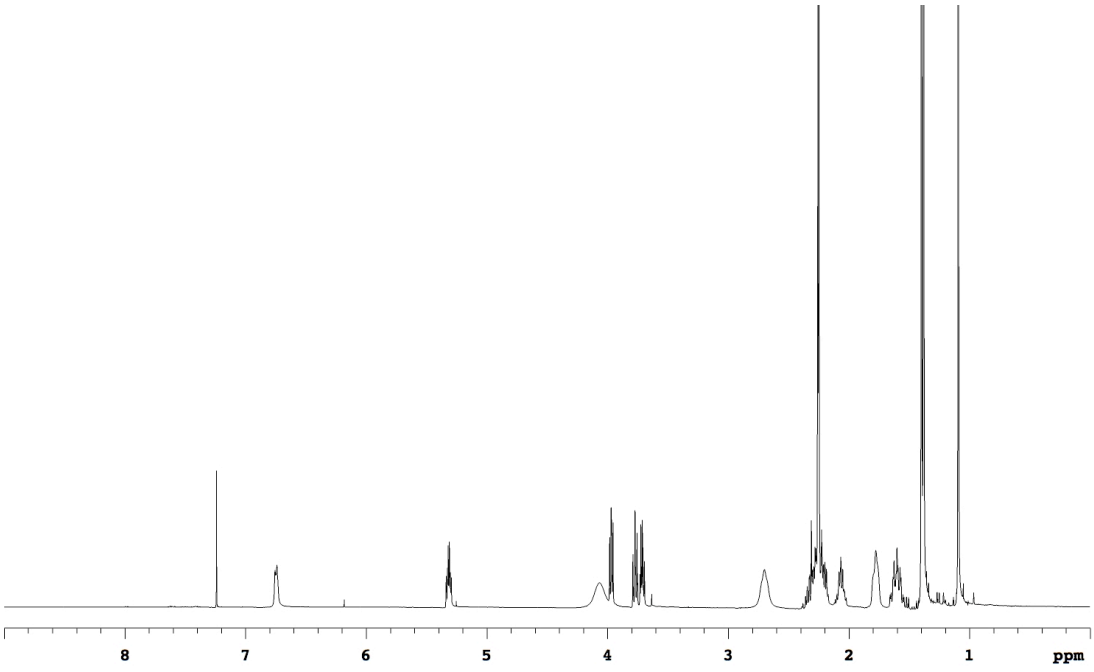
12s

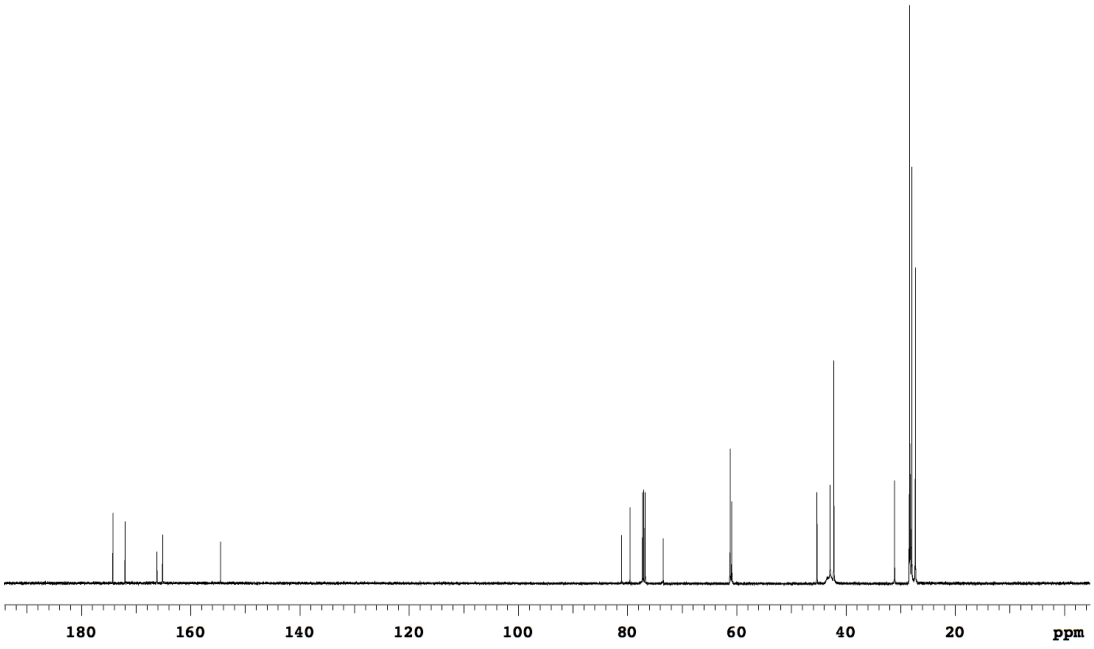
Compound **12s** was prepared from (0.44 g, 1.2 mmol) of **11s**. Flash chromatography (5 % to 10 % MeOH/CH_2Cl_2) afforded 0.62 g (89 %) **12s** as a white solid.

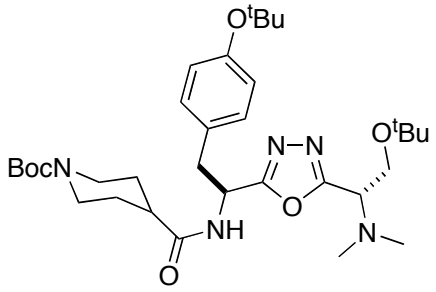
^{1}H NMR (500 MHz, $CDCl_3$) δ 6.74 (br, 1H), 5.32 (dt, 1H, J = 4.9, 8.2 Hz), 4.20-3.90 (m, 3H), 3.82-3.67 (m, 2H), 2.70 (br, 2H), 2.37-2.15 (m, 10H), 2.13-2.01 (m, 1H), 1.78 (br, 2H), 1.67-1.54 (m, 2H), 1.42-1.35 (m, 18H), 1.09 (s, 9H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 174.2, 172.0, 166.2, 165.1, 154.5, 81.1, 79.5, 73.4, 61.2, 60.9, 45.3, 42.9, 42.2, 42.2, 31.1, 28.4, 28.3, 28.2, 28.0, 27.3

MS (MALDI, m/z) 582.0 $(M+H)^+$


8
7
6
5
4
3
2
1
ppm


180
160
140
120
100
80
60
40
20
ppm



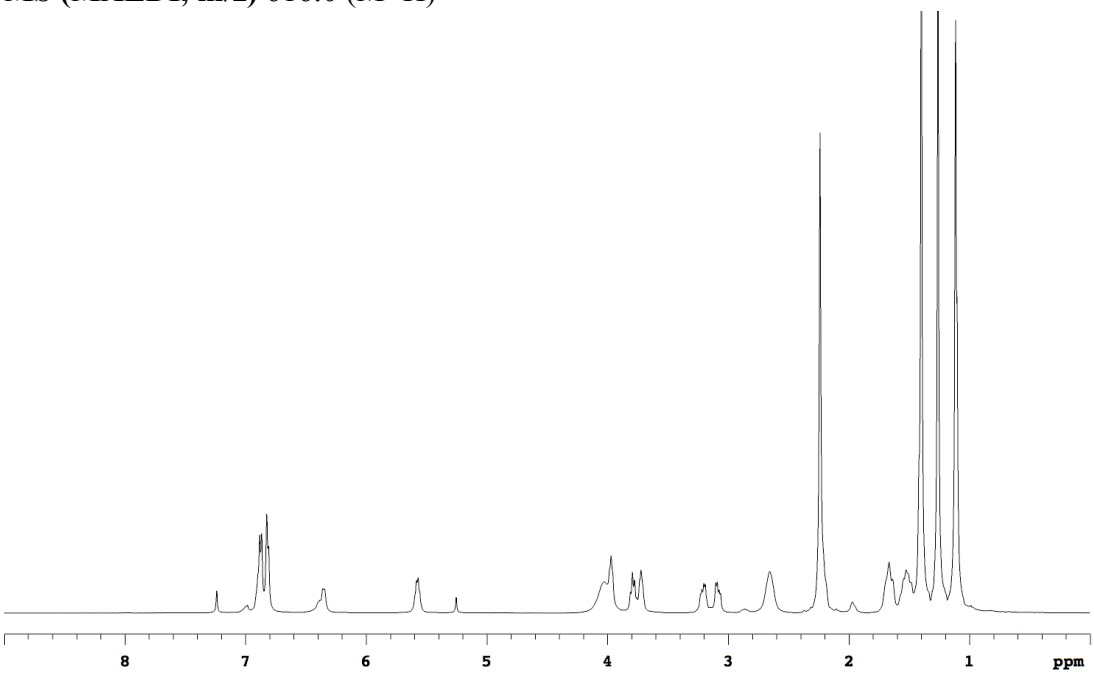
12t

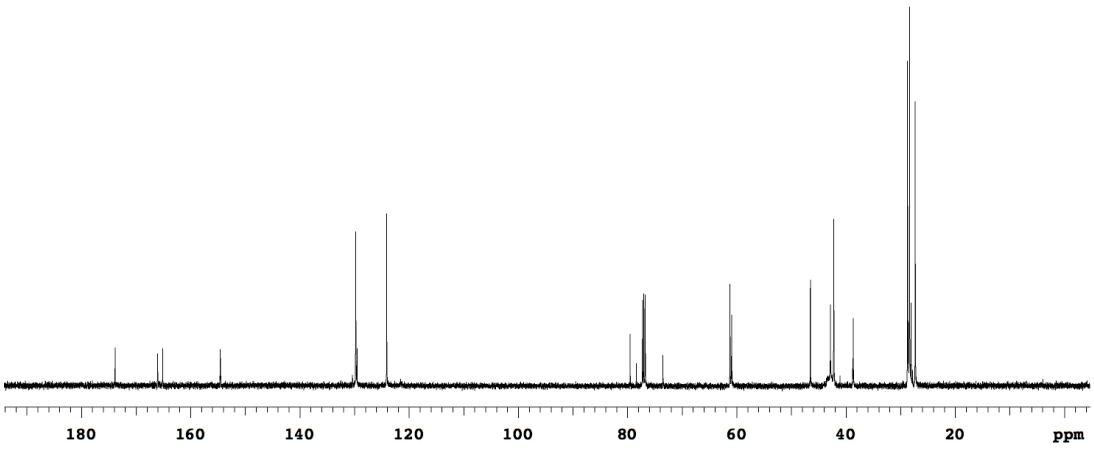
Compound **12t** was prepared from (1.54 g, 3.8 mmol) of **11t**. Flash chromatography (2 % to 3 % MeOH/CH_2Cl_2) afforded 1.65 g (71 %) **12t** as a white solid.

^{1}H NMR (500 MHz, $CDCl_3$) δ 6.96-6.75 (m, 4H), 6.36 (br, 1H), 5.65-5.51 (m, 1H), 4.20-3.88 (m, 3H), 3.86-3.66 (m, 2H), 3.21 (dd, 1H, J = 5.8, 13.6 Hz), 3.08 (dd, 1H, J = 5.7, 13.6 Hz), 2.66 (br, 2H), 2.30-2.16 (m, 7H), 1.75-1.61 (m, 2H), 1.60-1.46 (m, 2H), 1.40 (s, 9H), 1.26 (s, 9H), 1.12 (s, 9H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 173.8, 166.0, 165.1, 154.6, 154.5, 129.8, 129.5, 124.1, 79.5, 78.4, 73.5, 61.2, 61.0, 60.9, 46.5, 42.8, 42.2, 38.7, 38.6, 28.7, 28.3, 27.3

MS (MALDI, m/z) 616.0 $(M+H)^+$

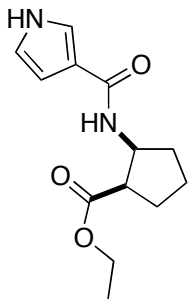



180
160
140
120
100
80
60
40
20
ppm

APPENDIX D

EXPERIMENTAL FOR CHAPTER V

General Methods. All chemicals were obtained from commercial suppliers and used without further purification. Dichloromethane was obtained anhydrous by distillation over calcium hydride and THF was distilled over sodium metal and benzophenone. Analytical HPLC analyses were carried out on 25 x 0.46 cm C-18 column using gradient conditions (5 – 95% B). The eluents used were: solvent A (H_2O with 0.1% TFA) and solvent B (CH_3CN with 0.1% TFA). The flow rate used was 1.0 mL/min. NMR spectra were recorded at 300 MHz and 500 MHz. NMR chemical shifts were expressed in ppm relative to internal solvent peaks, and coupling constants were measured in Hz.

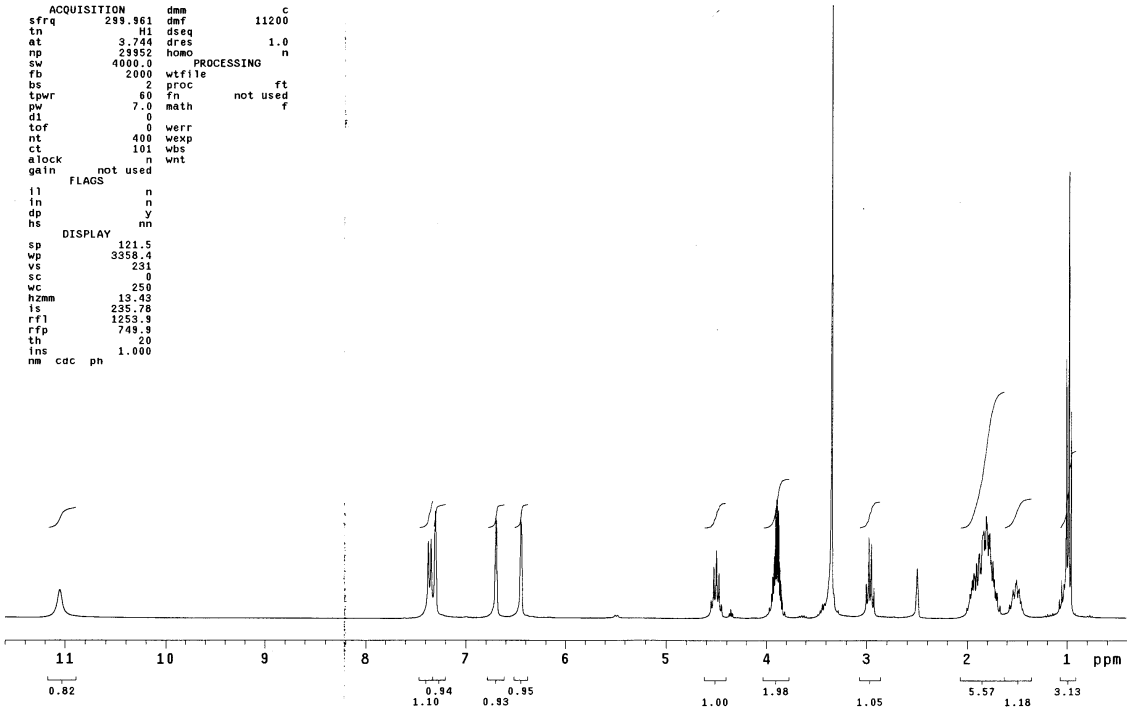


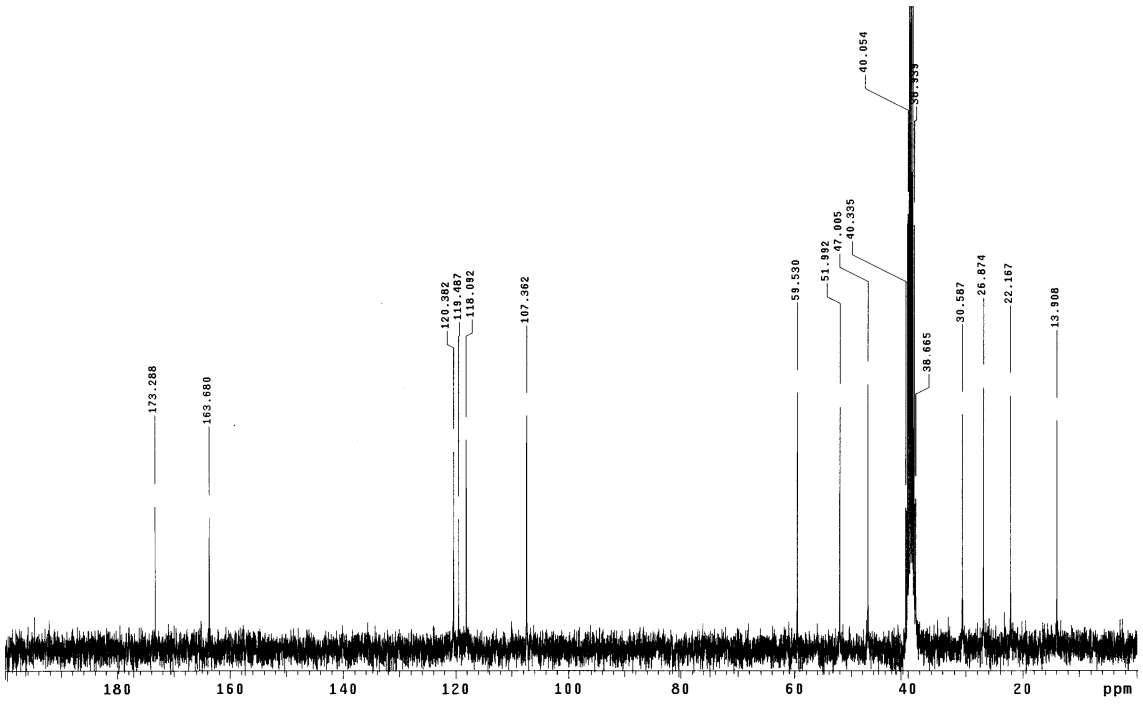
Preparation of 20. To a solution of 1*H*-pyrrole-3-carboxylic acid (528.5 mg, 4.8 mmol) in CH_2Cl_2 (25 mL) and DMF (10 mL) mixture, ethyl cis-2-amino-1-cyclopentane carboxylate hydrochloride (774.7 mg, 4.0 mmol) was added at 0 °C. Then 4-(dimethylamino)pyridine (977.4 mg, 8.0 mmol) and *N*, *N*′-diisopropylcarbodiimide (0.94 mL, 6.0 mmol) was added to the above solution. The reaction mixture was stirred at 25 °C for 10 h. The mixture was washed with saturated $NaHCO_3$ (4 × 30 mL) and brine (30 mL). The organic layer was separated and dried over Na_2SO_4 and concentrated to dryness. Recrystallization from ethanol yielded pure compound **20 (**650 mg, 65 %) as a beige amorphous powder.

^{1}H NMR (300 MHz, DMSO-d_6) δ 11.06 (br, 1H), 7.42-7.27 (m, 2H), 6.74-6.67 (m, 1H), 6.50-6.42 (m, 1H), 4.59-4.43 (m, 1H), 4.00-3.80 (m, 2H), 3.03-2.90 (m, 1H), 2.02-1.66 (m, 5H), 1.60-1.42 (m, 1H), 0.99 (t, 3H, J = 7.0 Hz)

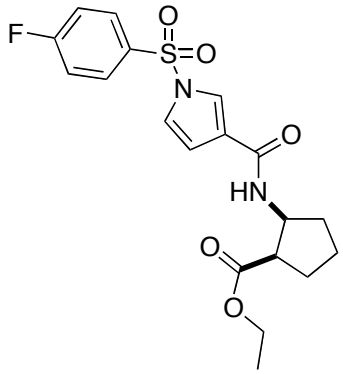
^{13}C NMR (75 MHz, DMSO-d_6) δ 173.3, 163.7, 120.4, 119.5, 118.1, 107.4, 59.5, 52.0, 47.0, 30.6, 26.9, 22.2, 13.9

MS (ESI, m/z) 251.1 $(M+H)^+$





General Procedure for Preparation of 22a-k. Pyrrole **20** (1 equiv) was added to a well-agitated suspension of NaOH (3 equiv) in dichloroethane (0.10 M). This mixture was then cooled to 0 °C and stirred for 20 min, following which a solution of sulfonyl chloride (1.2 equiv) in dichloroethane (0.24 M) was added dropwise over a period of 15 min. Thirty minutes after the completion of addition, the reaction was allowed to come to room temperature and left stirring overnight. After quenching the reaction by pouring onto 10 mL of H_2O, the organic layer was separated, and the aqueous layer was extracted with CH_2Cl_2 (3 × 10 mL). The combined organic extract was washed with H_2O to neutrality and dried over Na_2SO_4. Removal of the solvent *in vacuo* gave crude product. Flash chromatography with EtOAc/hexanes mixture provided pure **22a-k**.



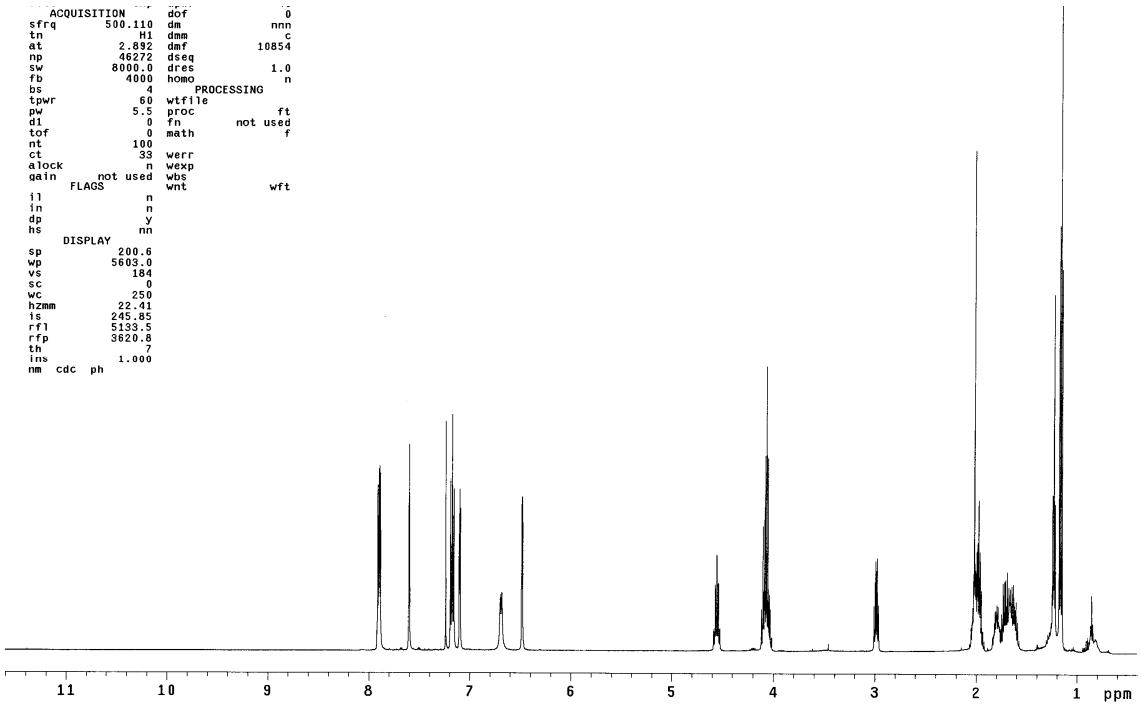
22a

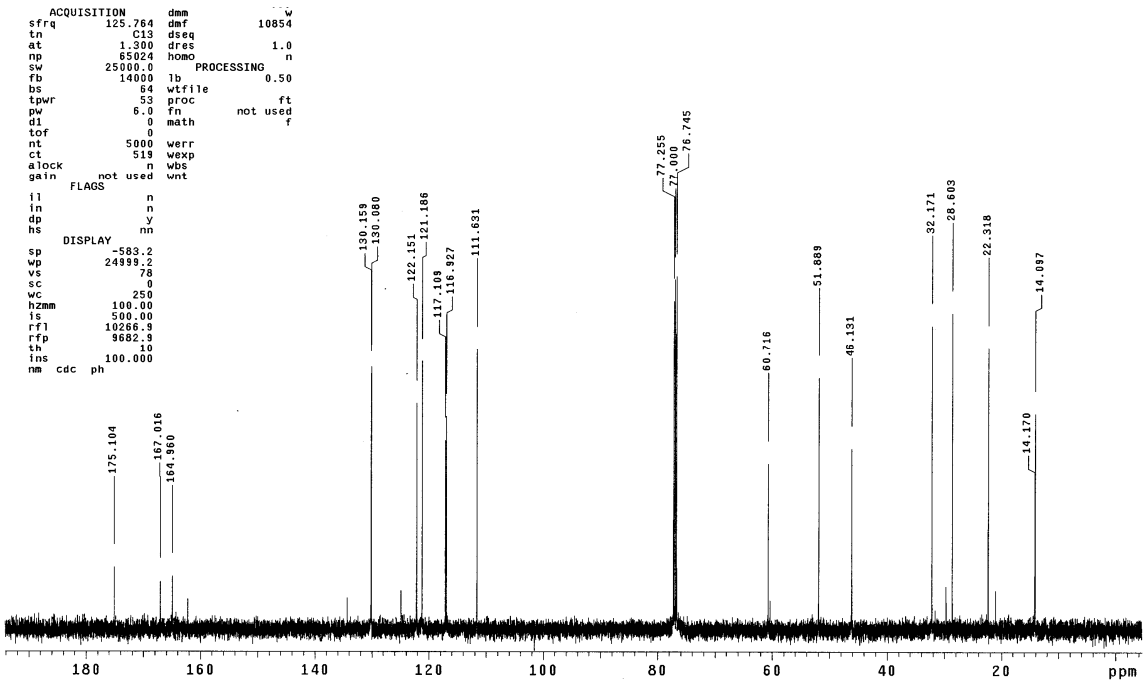
Compound **22a** was prepared from (25 mg, 0.1 mmol) of **20** and 4-fluorobenzenesulfonyl chloride. Flash chromatography (15 % to 25 % EtOAc/Hexanes) afforded 17.7 mg (43 %) **22a** as a white solid.

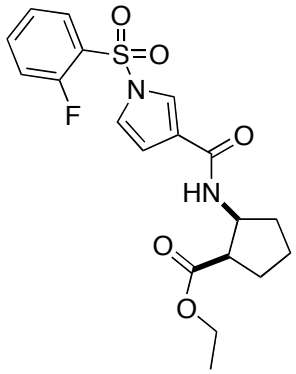
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.92-7.87 (m, 2H), 7.62-7.58 (m, 1H), 7.21-7.14 (m, 2H), 7.12-7.08 (m, 1H), 6.69 (d, 1H, J = 8.0 Hz), 6.50-6.46 (m, 1H), 4.60-4.53 (m, 1H), 4.13-4.01 (m, 2H), 3.02-2.95 (m, 1H), 2.06-1.92 (m, 3H), 1.84-1.57 (m, 3H), 1.17 (t, 3H, J = 7.0 Hz)

^{13}C NMR (125 MHz, $CDCl_3$) δ 175.1, 167.0, 165.0, 130.2, 130.1, 122.2, 121.2, 117.1, 116.9, 111.6, 60.7, 51.9, 46.1, 32.1, 28.6, 22.3, 14.1

MS (ESI, m/z) 409.1 $(M+H)^+$


ACQUISITION
sfrq 500.110
tn H1
at 2.892
np 46272
sw 8000.0
fb 4000
bs 4
tpwr 60
pw 5.5
d1 0
tof 0
nt 100
ct 33
alock n
gain not used
FLAGS
il n
in n
dp y
hs nn
DISPLAY
sp 200.6
wp 5603.0
vs 184
sc 0
wc 250
hzmm 22.41
is 245.85
rfl 5133.5
rfp 3620.8
th 7
ins 1.000
nm cdc ph
dof 0
dm nnn
dmm c
dmf 10854
dseq
dres 1.0
homo n
PROCESSING
wtfile
proc ft
fn not used
math f
werr
wexp
wbs
wnt wft
11 10 9 8 7 6 5 4 3 2 1 ppm


ACQUISITION
sfrq 125.764
tn C13
at 1.300
np 65024
sw 25000.0
fb 14000
bs 64
tpwr 53
pw 6.0
d1 0
tof 0
nt 5000
ct 519
alock n
gain not used
FLAGS
il n
in n
dp y
hs nn
DISPLAY
sp -583.2
wp 24999.2
vs 78
sc 0
wc 250
hzmm 100.00
is 500.00
rfl 10266.9
rfp 9682.9
th 10
ins 100.000
nm cdc ph
dmm w
dmf 10854
dseq
dres 1.0
homo n
PROCESSING
lb 0.50
wtfile
proc ft
fn not used
math f
werr
wexp
wbs
wnt
175.104
167.016
164.360
130.159
130.080
122.151
121.186
117.109
116.927
111.631
77.255
77.000
76.745
60.716
51.889
48.131
32.171
28.603
22.318
14.087
14.170
180 160 140 120 100 80 60 40 20 ppm



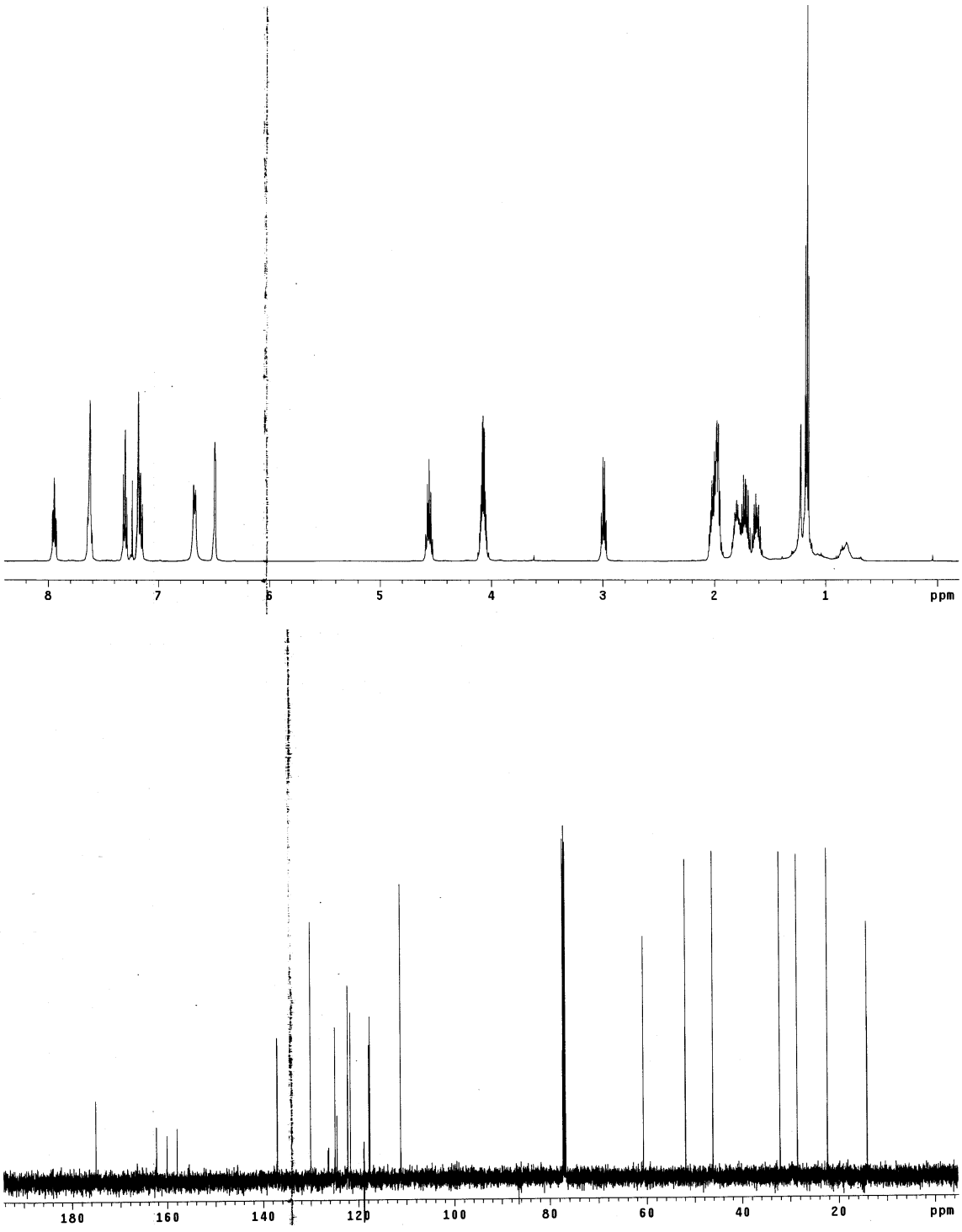
22b

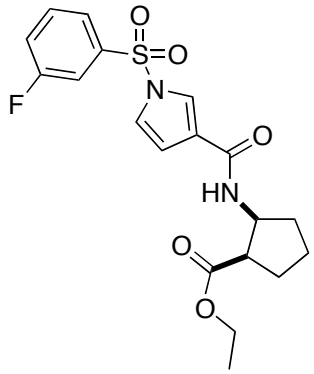
Compound **22b** was prepared from (25 mg, 0.1 mmol) of **20** and 2-fluorobenzenesulfonyl chloride. Flash chromatography (15 % to 25 % EtOAc/Hexanes) afforded 28.3 mg (69 %) **22b** as a white solid.

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.98-7.92 (m, 1H), 7.62 (br, 2H), 7.34-7.28 (m, 1H), 7.22-7.13 (m, 2H), 6.68 (d, 1H, J = 8.0 Hz), 6.52-6.48 (m, 1H), 4.61-4.52 (m, 1H), 4.13-4.02 (m, 2H), 3.04-2.96 (m, 1H), 2.06-1.92 (m, 3H), 1.85-1.56 (m, 3H), 1.17 (t, 3H, J = 7.2 Hz)

^{13}C NMR (125 MHz, $CDCl_3$) δ 175.0, 162.3, 160.0, 158.0, 137.0, 130.0, 124.0, 122.2, 121.6, 117.8, 117.6, 111.2, 60.7, 51.9, 46.1, 32.1, 28.5, 22.3, 14.1

MS (ESI, m/z) 409.1 $(M+H)^+$


8
7
6
5
4
3
2
1
ppm
180
160
140
120
100
80
60
40
20
ppm



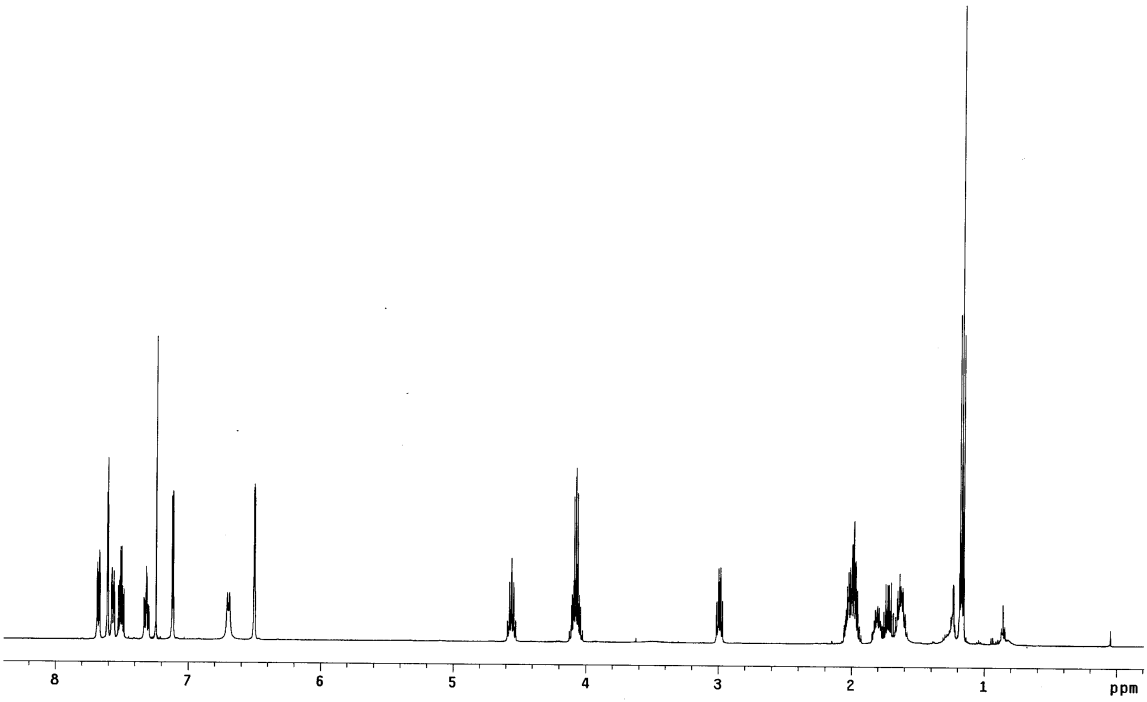
22c

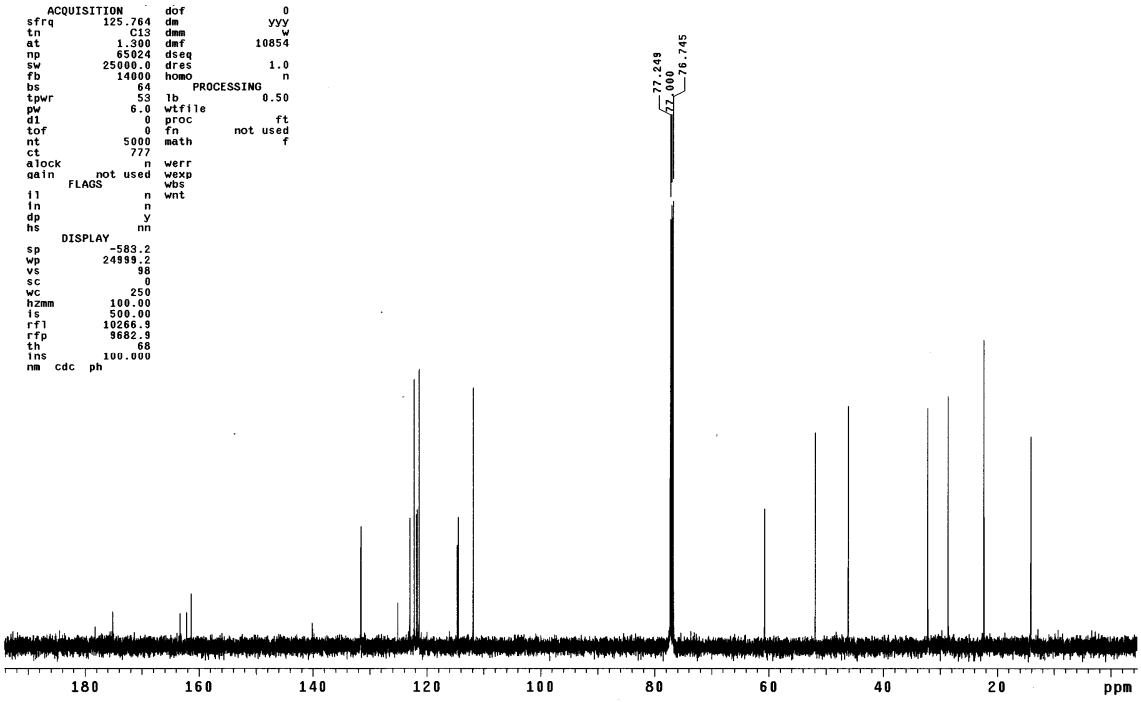
Compound **22c** was prepared from (25 mg, 0.1 mmol) of **20** and 3-fluorobenzenesulfonyl chloride. Flash chromatography (15 % to 25 % EtOAc/Hexanes) afforded 18.2 mg (45 %) **22c** as a white solid.

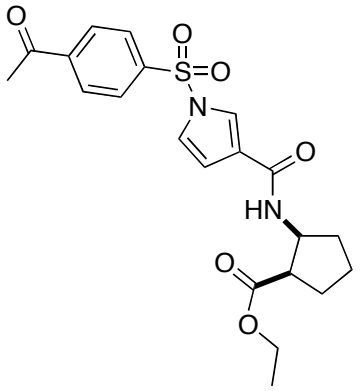
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.67 (dq, 1H, J = 8.0, 1.0 Hz), 7.62-7.59 (m, 1H), 7.56 (dt, 1H, J = 8.0, 2.0 Hz), 7.50 (td, 1H, J = 8.0, 5.0 Hz), 7.34-7.29 (m, 1H), 7.13-7.10 (m, 1H), 6.70 (d, 1H, J = 8.0 Hz), 6.52-6.49 (m, 1H), 4.60-4.52 (m, 1H), 4.13-4.02 (m, 2H), 3.02-2.96 (m, 1H), 2.07-1.92 (m, 3H), 1.85-1.57 (m, 3H), 1.17 (t, 3H, J = 7.0 Hz)

^{13}C NMR (125 MHz, $CDCl_3$) δ 175.1, 163.3, 162.2, 161.3, 140.0, 131.5, 125.0, 123.0, 122.2, 121.3, 114.5, 111.8, 60.8, 51.9, 46.1, 32.1, 28.6, 22.3, 14.1

MS (ESI, m/z) 409.1$(M+H)^+$


8
7
6
5
4
3
2
1
ppm


ACQUISITION
sfrq 125.764
tn C13
at 1.300
np 65024
sw 25000.0
fb 14000
bs 64
tpwr 53
pw 6.0
d1 0
tof 0
nt 5000
ct 777
alock n
gain not used
FLAGS
il n
in n
dp y
hs nn
DISPLAY
sp -583.2
wp 24999.2
vs 98
sc 0
wc 250
hzmm 100.00
is 500.00
rfl 10266.9
rfp 9682.9
th 68
ins 100.000
nm cdc ph
dof 0
dm yyy
dmm w
dmf 10854
dseq
dres 1.0
homo n
PROCESSING
lb 0.50
wtfile
proc ft
fn not used
math f
werr
wexp
wbs
wnt
77.249
77.000
76.745
180
160
140
120
100
80
60
40
20
ppm



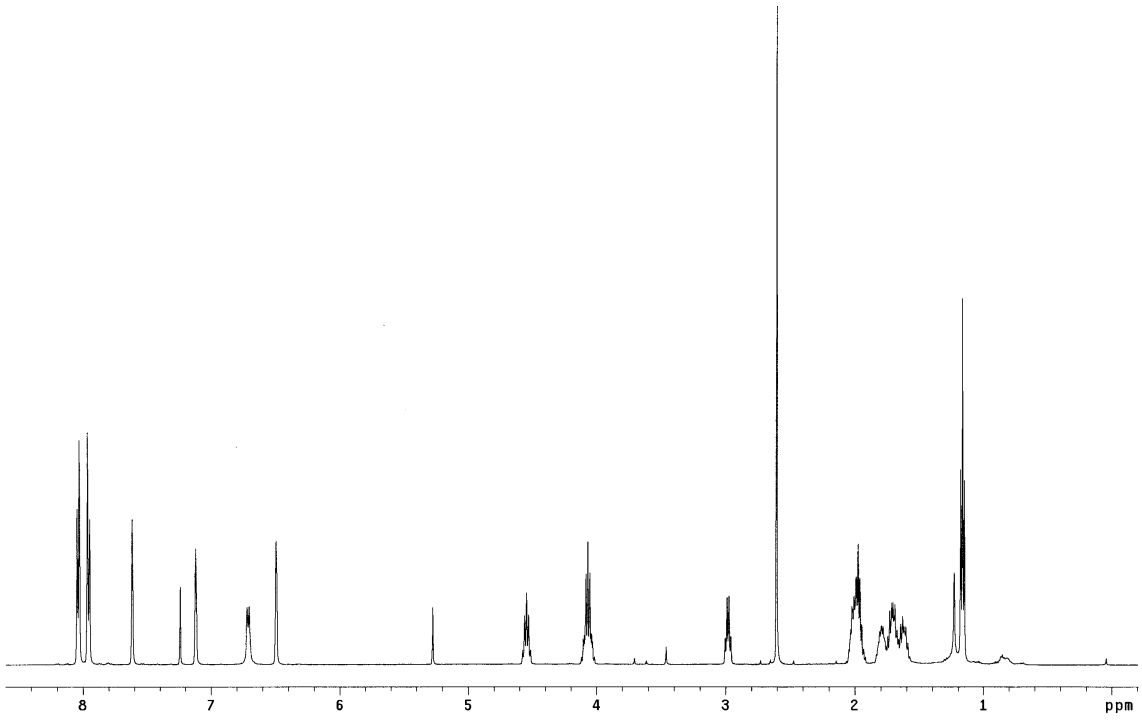
22d

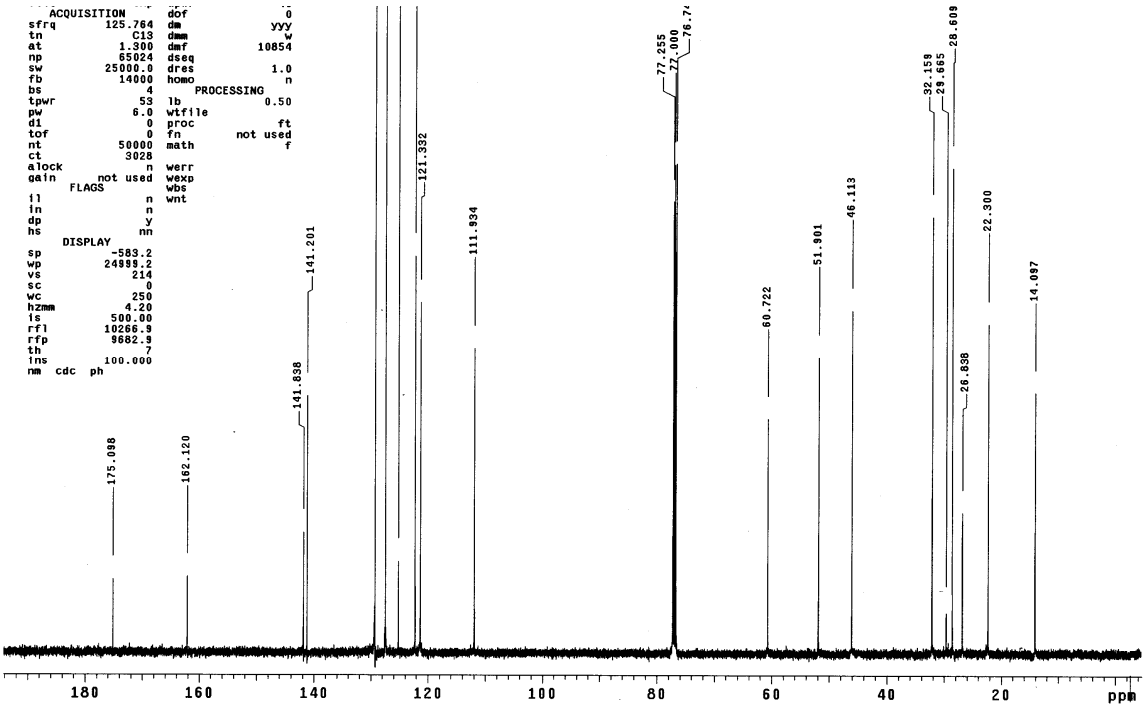
Compound **22d** was prepared from (25 mg, 0.1 mmol) of **20** and 4-acetylbenzenesulfonyl chloride. Flash chromatography (15 % to 25 % EtOAc/Hexanes) afforded 18.9 mg (44 %) **22d** as a beige solid.

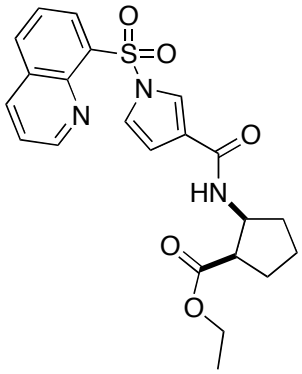
^{1}H NMR (500 MHz, $CDCl_3$) δ 8.04 (d, 2H, J = 8.0 Hz), 7.95 (d, 2H, J = 8.0 Hz), 7.62 (br, 1H), 7.12 (br, 1H), 6.71 (d, 1H, J = 8.0 Hz), 6.50 (br, 1H), 4.59-4.50 (m, 1H), 4.13-4.01 (m, 2H), 3.02-2.94 (m, 1H), 2.60 (s, 3H), 2.06-1.91 (m, 3H), 1.84-1.57 (m, 3H), 1.17 (t, 3H, J = 7.2 Hz)

^{13}C NMR (125 MHz, $CDCl_3$) δ 175.1, 162.1, 141.8, 141.2, 129.3, 129.3, 127.4, 125.2, 122.2, 121.3, 111.9, 60.7, 51.9, 46.1, 32.2, 28.6, 26.8, 22.3, 14.1

MS (ESI, m/z) 455.1 $(M+Na)^+$


8
7
6
5
4
3
2
1
ppm


ACQUISITION
sfrq 125.764
tn C13
at 1.300
np 65024
sw 25000.0
fb 14000
bs 4
tpwr 53
pw 6.0
d1 0
tof 0
nt 50000
ct 3028
alock n
gain not used
FLAGS
il n
in n
dp y
hs nn
DISPLAY
sp -583.2
wp 24995.2
vs 214
sc 0
wc 250
hzmm 4.20
is 500.00
rfl 10266.9
rfp 9682.9
th 7
ins 100.000
nm cdc ph
dof 0
dm yyy
dmm w
dmf 10854
dseq
dres 1.0
homo n
PROCESSING
lb 0.50
wtfile
proc ft
fn not used
math f
werr
wexp
wbs
wnt
175.098
162.120
141.838
141.201
121.332
111.934
77.255
77.000
60.722
51.901
46.113
32.158
29.665
28.609
26.838
22.300
14.097
180
160
140
120
100
80
60
40
20
ppm



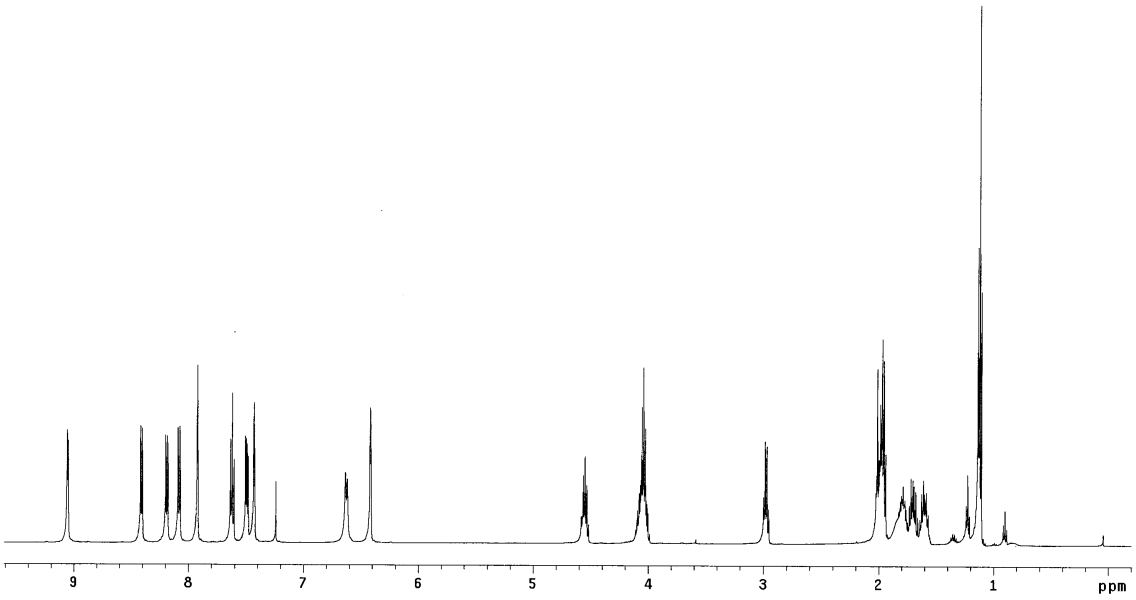
22e

Compound **22e** was prepared from (25 mg, 0.1 mmol) of **20** and 8-quinolinebenzenesulfonyl chloride. Flash chromatography (20 % to 70 % EtOAc/Hexanes) afforded 31.1 mg (71 %) **22e** as a beige solid.

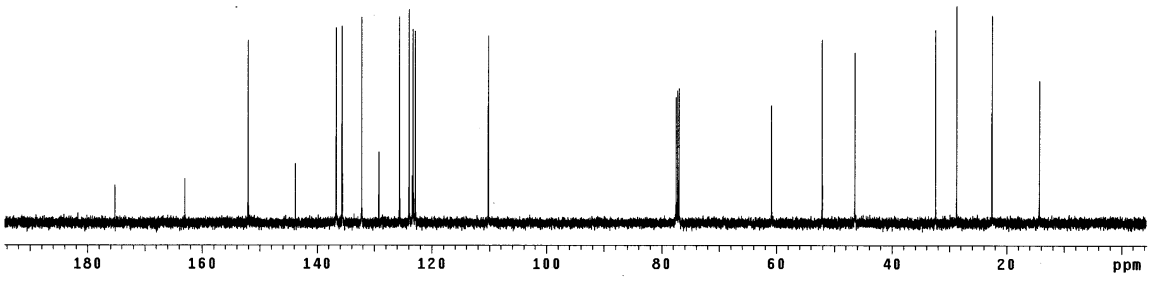
^{1}H NMR (500 MHz, $CDCl_3$) δ 9.10-9.02 (m, 1H), 8.41 (d, 1H, J = 8.0 Hz), 8.19 (d, 1H, J = 8.0 Hz), 8.09 (d, 1H, J = 8.0 Hz), 7.97-7.90 (m, 1H), 7.68-7.59 (m, 1H), 7.55-7.47 (m, 1H), 7.47-7.39 (m, 1H), 6.63 (d, 1H, J = 8.0 Hz), 6.47-6.38 (m, 1H), 4.61-4.51 (m, 1H), 4.13-3.98 (m, 2H), 3.03-2.94 (m, 1H), 2.06-1.92 (m, 3H), 1.90-1.54 (m, 3H), 1.13 (t, 3H, J = 7.6 Hz)

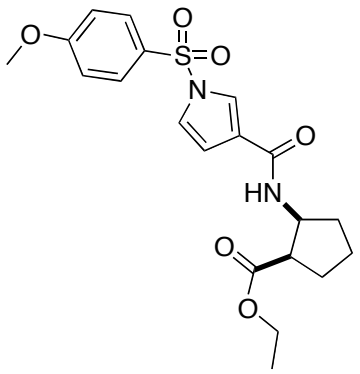
^{13}C NMR (125 MHz, $CDCl_3$) δ 175.2, 163.0, 152.0, 143.8, 136.7, 135.7, 135.6, 132.2, 129.2, 125.6, 123.9, 123.4, 123.2, 122.8, 110.2, 60.9, 52.1, 46.5, 32.4, 28.7, 22.5, 14.0

MS (ESI, m/z) 464.1 $(M+Na)^+$



ACQUISITION			
sfrq	125.764	dof	0
tn	C13	dm	yyy
at	1.300	dmm	w
np	65024	dmf	10854
sw	25000.0	dseq	
fb	14000	dres	1.0
bs	64	homo	n
tpwr	53	PROCESSING	
pw	6.0	lb	0.50
d1	0	wtfile	
tof	0	proc	ft
nt	1000	fn	not used
ct	305	math	f
alock	n	werr	
gain	not used	wexp	
FLAGS		wbs	
il	n	wnt	
in	n		
dp	y		
hs	nn		
DISPLAY			
sp	-555.0		
wp	24999.2		
vs	47		
sc	0		
wc	250		
hzmm	100.00		
is	500.00		
rfl	10238.6		
rfp	9682.9		
th	68		
ins	100.000		
nm cdc ph			





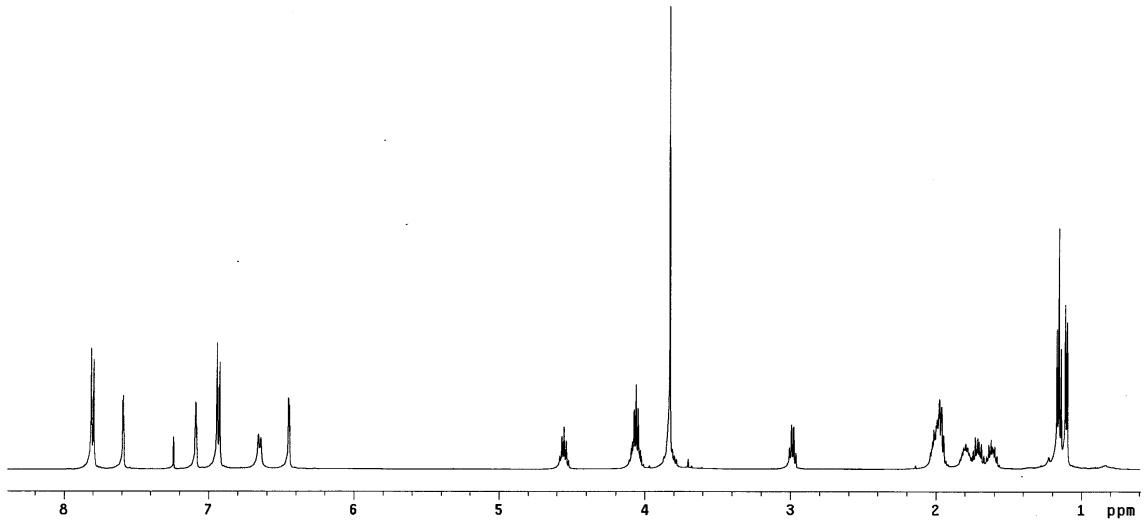
22f

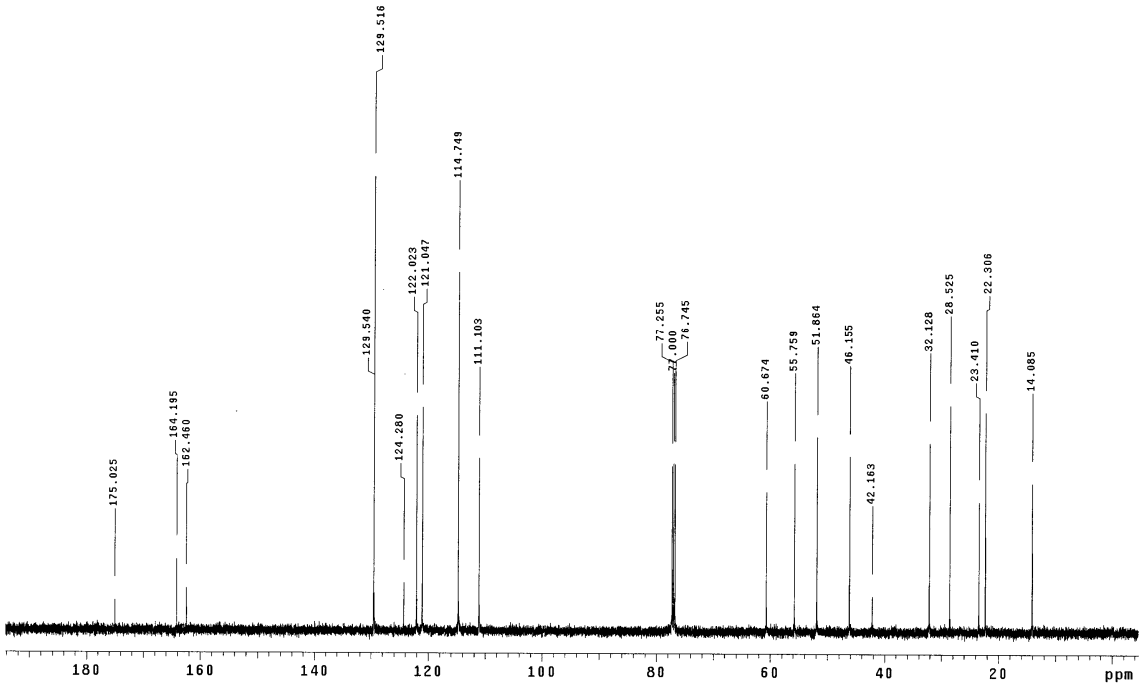
Compound **22f** was prepared from (25 mg, 0.1 mmol) of **20** and 4-methoxybenzenesulfonyl chloride. Flash chromatography (20 % to 30 % EtOAc/Hexanes) afforded 37.6 mg (90 %) **22f** as a white solid.

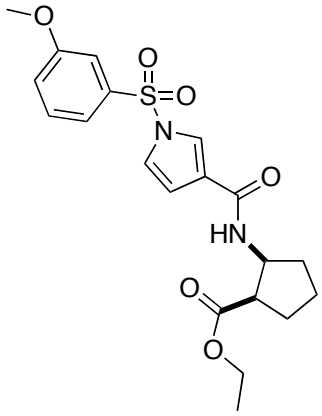
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.80 (d, 2H, J = 9.0 Hz), 7.59 (br, 1H), 7.12-7.07 (m, 1H), 6.93 (d, 2H, J = 9.0 Hz), 6.65 (d, 1H, J = 8.8 Hz), 6.48-6.43 (m, 1H), 4.60-4.51 (m, 1H), 4.12-4.00 (m, 2H), 3.83 (s, 3H), 3.03-2.95 (m, 1H), 2.06-1.91 (m, 3H), 1.87-1.58 (m, 3H), 1.15 (t, 3H, J = 6.7 Hz)

^{13}C NMR (125 MHz, $CDCl_3$) δ 175.0, 164.2, 162.5, 129.5, 129.5, 124.3, 122.0, 121.0, 114.7, 111.1, 60.8, 55.8, 51.9, 46.2, 32.1, 28.5, 22.3, 14.1

MS (ESI, m/z) 421.1 $(M+H)^+$


8
7
6
5
4
3
2
1
ppm


175.025
164.185
162.460
129.516
129.540
124.280
122.023
121.047
114.749
111.103
77.255
77.000
76.745
60.674
55.759
51.864
46.155
42.163
32.128
28.525
23.410
22.306
14.085
180
160
140
120
100
80
60
40
20
ppm



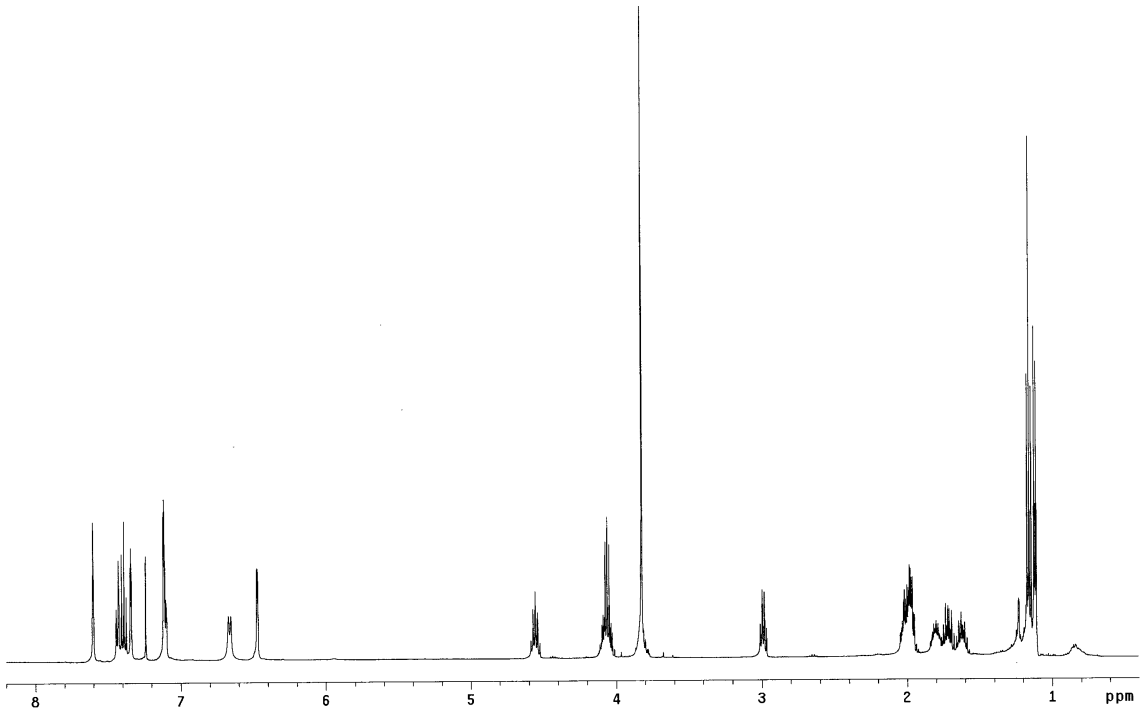
22g

Compound **22g** was prepared from (25 mg, 0.1 mmol) of **20** and 3-methoxybenzenesulfonyl chloride. Flash chromatography (20 % to 30 % EtOAc/Hexanes) afforded 19.0 mg (45 %) **22g** as a white solid.

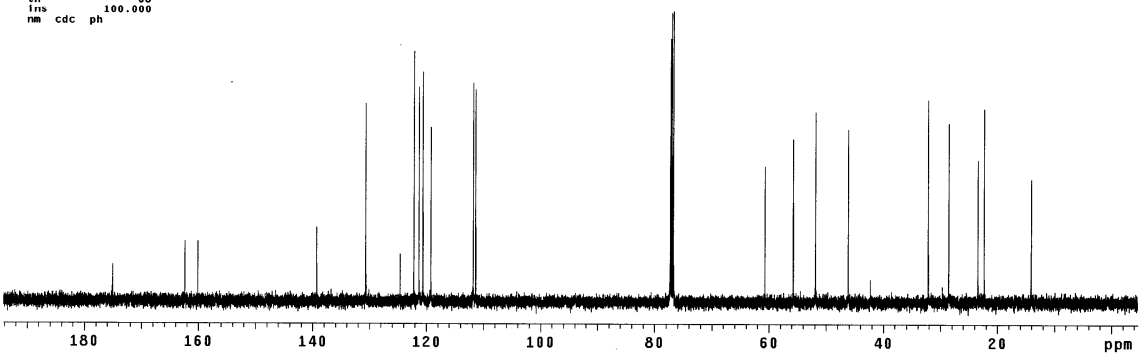
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.62-7.59 (m, 1H), 7.46-7.37 (m, 2H), 7.36-7.33 (m, 1H), 7.14-7.08 (m, 2H), 6.66 (d, 1H, J = 8.6 Hz), 6.50-6.48 (m, 1H), 4.60-4.51 (m, 1H), 4.13-4.01 (m, 2H), 3.82 (s, 3H), 3.03-2.96 (m, 1H), 2.06-1.92 (m, 3H), 1.87-1.57 (m, 3H), 1.16 (t, 3H, J = 7.4 Hz)

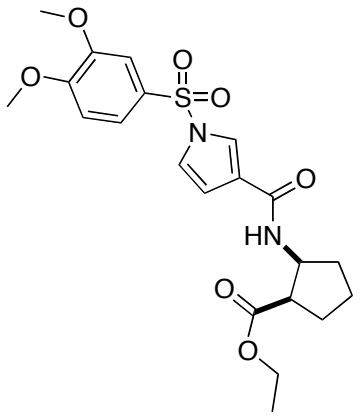
^{13}C NMR (125 MHz, $CDCl_3$) δ 175.1, 162.4, 160.1, 139.3, 130.7, 124.6, 122.2, 121.3, 120.6, 119.2, 111.8, 111.4, 60.7, 55.8, 51.9, 46.1, 32.2, 28.6, 22.3, 14.1

MS (ESI, m/z) 421.1 $(M+H)^+$



ACQUISITION			
sfrq	125.764	dof	0
tn	C13	dm	yyy
at	1.300	dmm	w
np	65024	dmf	10854
sw	25000.0	dseq	
fb	14000	dres	1.0
bs	64	homo	n
tpwr	53	PROCESSING	
pw	6.0	lb	0.50
d1	0	wtfile	
tof	0	proc	ft
nt	1000	fn	not used
ct	775	math	f
alock	n	werr	
gain	not used	wexp	
FLAGS		wbs	
il	n	wnt	
in	n		
dp	y		
hs	nn		
DISPLAY			
sp	-584.0		
wp	24999.2		
vs	64		
sc	0		
wc	250		
hzmm	100.00		
is	500.00		
rfl	10267.6		
rfp	9682.9		
th	68		
ins	100.000		
nm cdc ph			





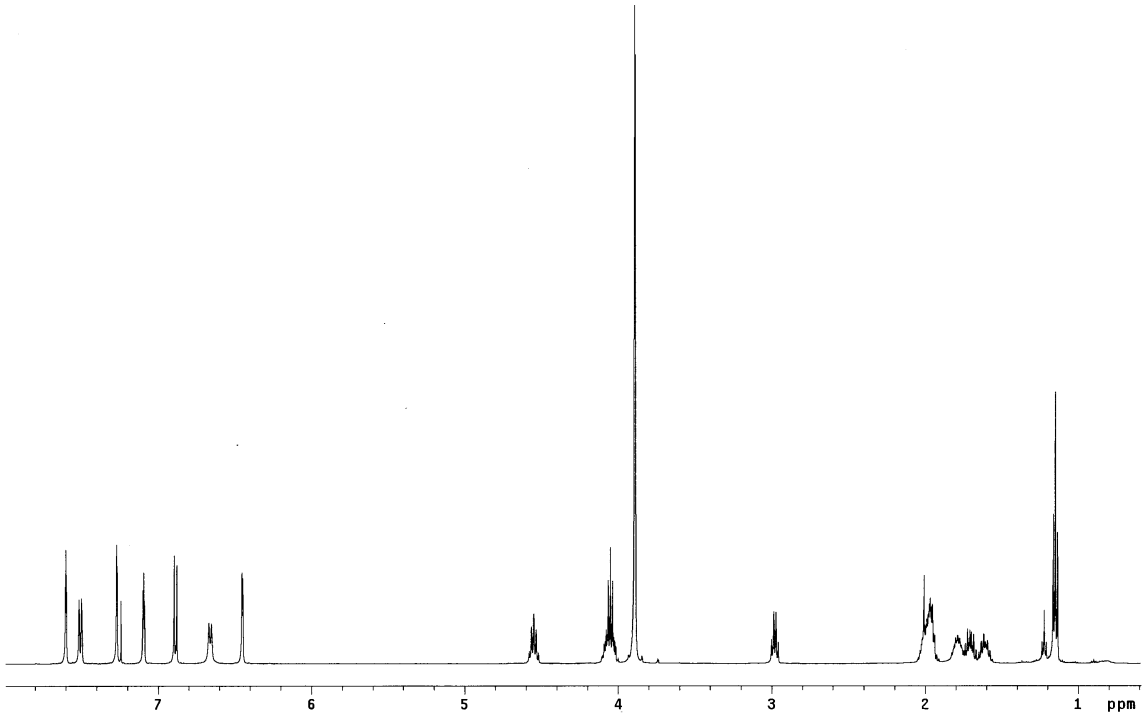
22h

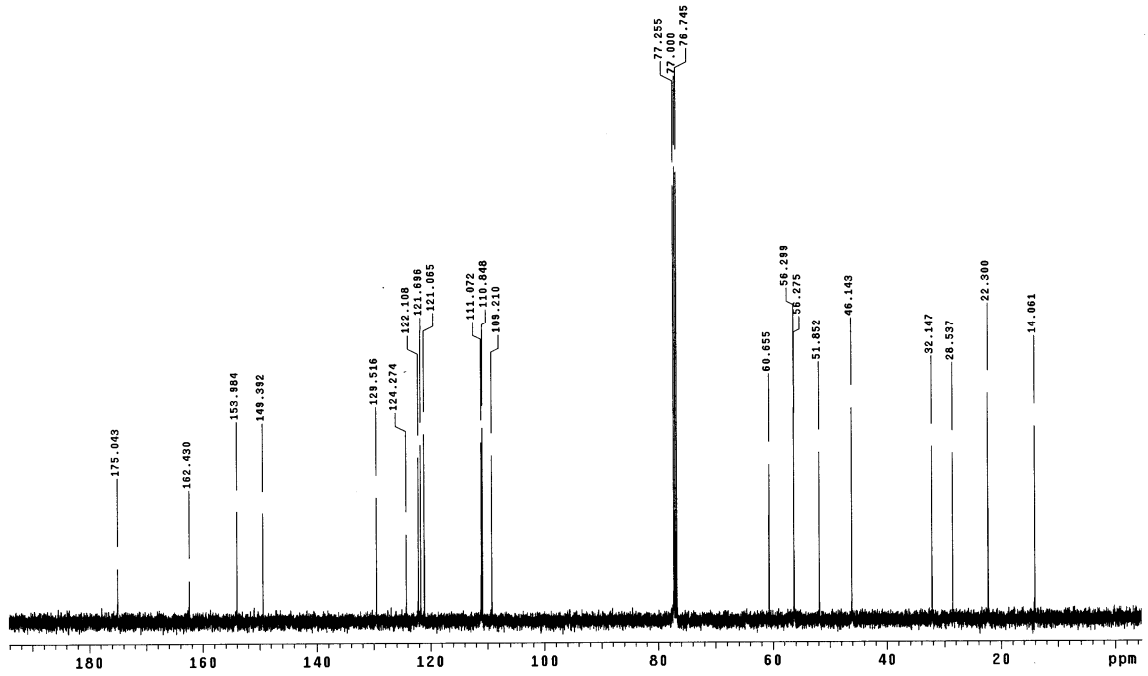
Compound **22h** was prepared from (25 mg, 0.1 mmol) of **20** and 3,4-dimethoxybenzenesulfonyl chloride. Flash chromatography (20 % to 35 % EtOAc/Hexanes) afforded 32.8 mg (73 %) **22h** as a white solid.

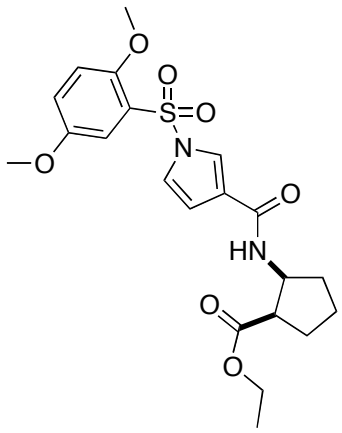
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.62-7.58 (m, 1H), 7.50 (dd, 1H, J = 8.4, 2.2 Hz), 7.29-7.26 (m, 1H), 7.12-7.08 (m, 1H), 6.89 (d, 1H, J = 9.5 Hz), 6.66 (d, 1H, J = 8.2 Hz), 6.47-6.43 (m, 1H), 4.59-4.51 (m, 1H), 4.11-4.00 (m, 2H), 3.89 (s, 3H), 3.88 (s, 3H), 3.02-2.95 (m, 1H), 2.06-1.90 (m, 3H), 1.86-1.55 (m, 3H), 1.17 (t, 3H, J = 7.5 Hz)

^{13}C NMR (125 MHz, $CDCl_3$) δ 175.0, 162.4, 154.0, 149.4, 129.5, 124.3, 122.1, 121.7, 121.1, 111.1, 110.9, 109.2, 60.7, 56.3, 56.3, 51.9, 46.1, 32.1, 28.5, 22.3, 14.1

MS (ESI, m/z) 451.2 $(M+H)^+$


7
6
5
4
3
2
1
ppm


175.043
162.430
153.984
149.392
129.516
124.274
122.108
121.698
121.065
111.072
110.648
109.210
77.255
77.000
76.745
60.655
56.299
56.275
51.852
46.143
32.147
28.537
22.300
14.061
180
160
140
120
100
80
60
40
20
ppm



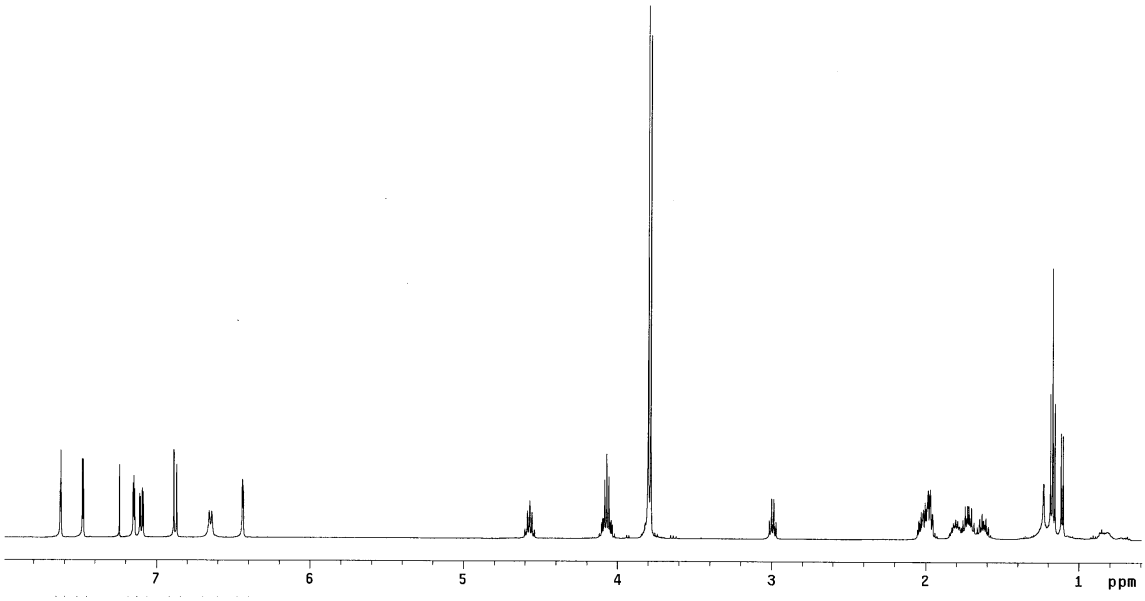
22i

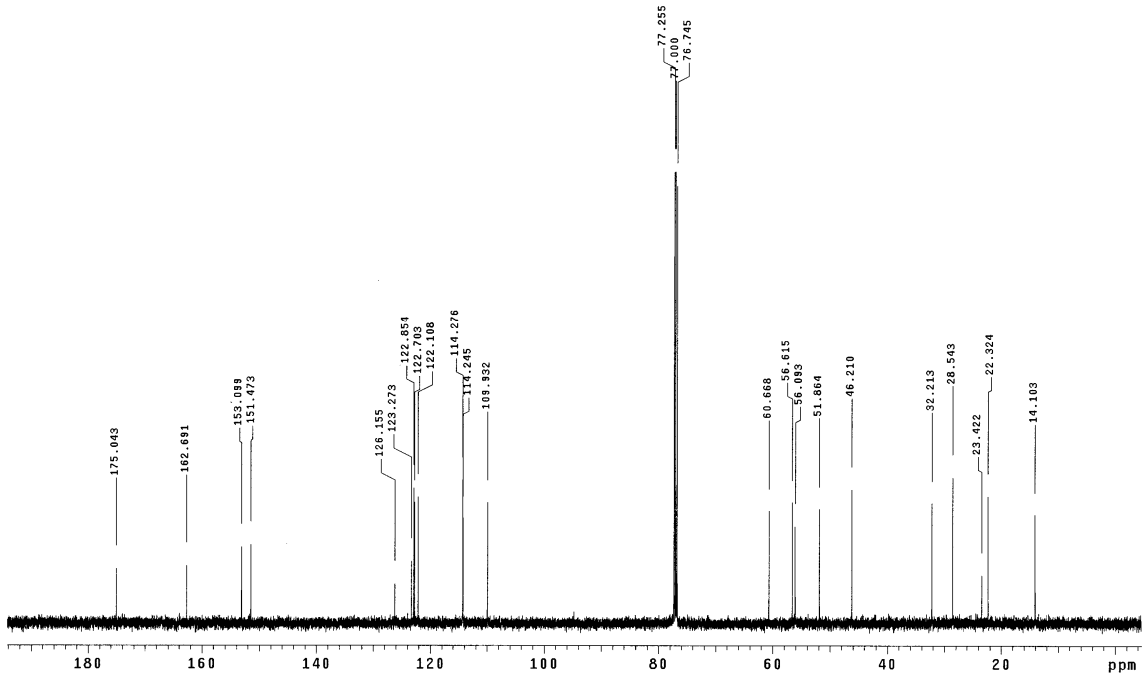
Compound **22i** was prepared from (25 mg, 0.1 mmol) of **20** and 2,5-dimethoxybenzenesulfonyl chloride. Flash chromatography (20 % to 35 % EtOAc/Hexanes) afforded 37.7 mg (84 %) **22i** as a white solid.

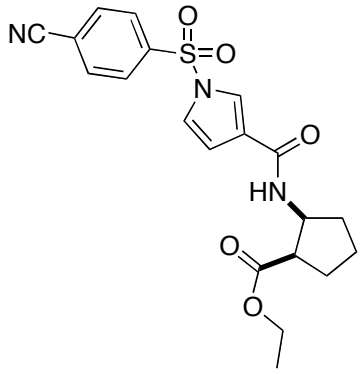
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.64-7.61 (m, 1H), 7.48 (d, 1H, J = 3.3 Hz), 7.17-7.13 (m, 1H), 7.10 (dd, 1H, J = 9.5, 3.3 Hz), 6.88 (d, 1H, J = 9.5 Hz), 6.65 (d, 1H, J = 8.3 Hz), 6.46-6.42 (m, 1H), 4.61-4.53 (m, 1H), 4.12-4.02 (m, 2H), 3.80 (s, 3H), 3.79 (s, 3H), 3.02-2.96 (m, 1H), 2.06-1.92 (m, 3H), 1.85-1.57 (m, 3H), 1.17 (t, 3H, J = 7.2 Hz)

^{13}C NMR (125 MHz, $CDCl_3$) δ 175.0, 162.7, 153.1, 151.5, 126.2, 123.3, 122.9, 122.7, 122.1, 114.3, 114.2, 109.9, 60.7, 56.6, 56.1, 51.9, 46.2, 32.2, 28.5, 22.3, 14.1

MS (ESI, m/z) 451.2 $(M+H)^+$


7
6
5
4
3
2
1
ppm


175.043
162.691
153.099
151.473
126.155
123.273
122.854
122.703
122.108
114.276
114.245
109.932
77.255
77.000
76.745
60.668
56.615
56.093
51.864
46.210
32.213
28.543
23.422
22.324
14.103
180
160
140
120
100
80
60
40
20
ppm



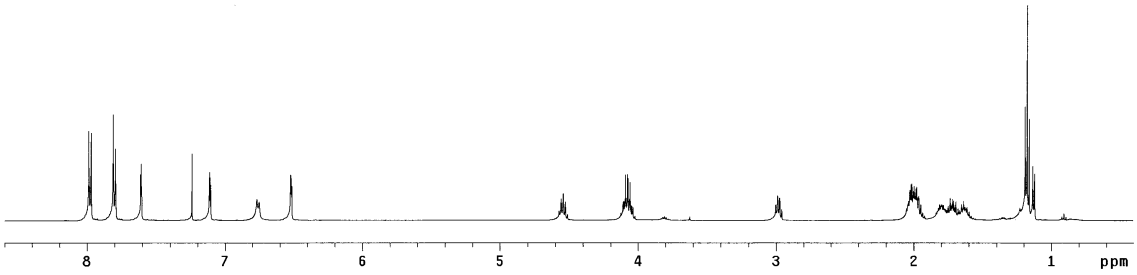
22j

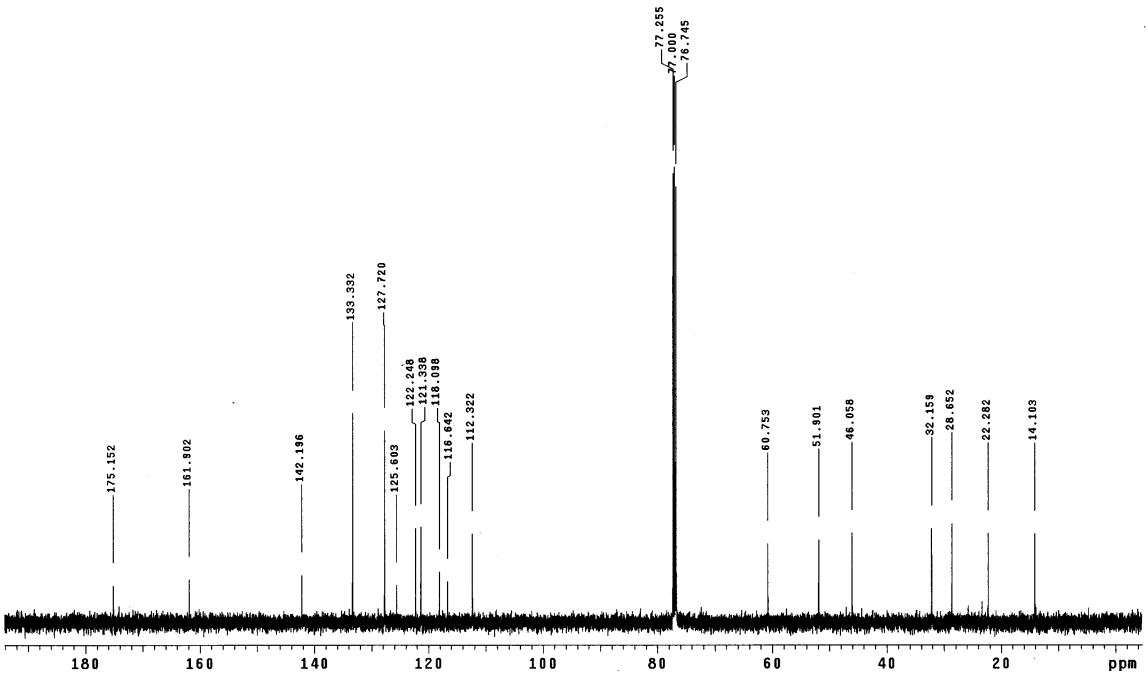
Compound **22j** was prepared from (25 mg, 0.1 mmol) of **20** and 4-cyanobenzenesulfonyl chloride. Flash chromatography (20 % to 30 % EtOAc/Hexanes) afforded 18.4 mg (44 %) **22j** as a white solid.

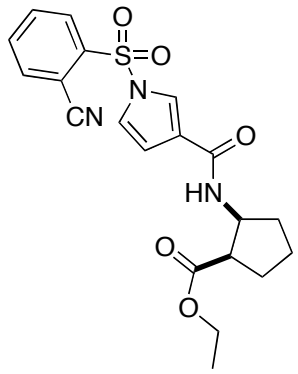
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.98 (d, 2H, J = 9.0 Hz), 7.80 (d, 2H, J = 9.0 Hz), 7.63-7.59 (m, 1H), 7.14-7.09 (m, 1H), 6.76 (d, 1H, J = 8.0 Hz), 6.55-6.49 (m, 1H), 4.61-4.50 (m, 1H), 4.16-4.02 (m, 2H), 3.04-2.95 (m, 1H), 2.09-1.55 (m, 6H), 1.18 (t, 3H, J = 7.3 Hz)

^{13}C NMR (125 MHz, $CDCl_3$) δ 175.2, 161.9, 142.2, 133.3, 127.7, 125.6, 122.2, 121.3, 118.1, 116.6, 112.3, 60.8, 51.9, 46.1, 32.2, 28.7, 22.3, 14.1

MS (ESI, m/z) 438.1 $(M+Na)^+$







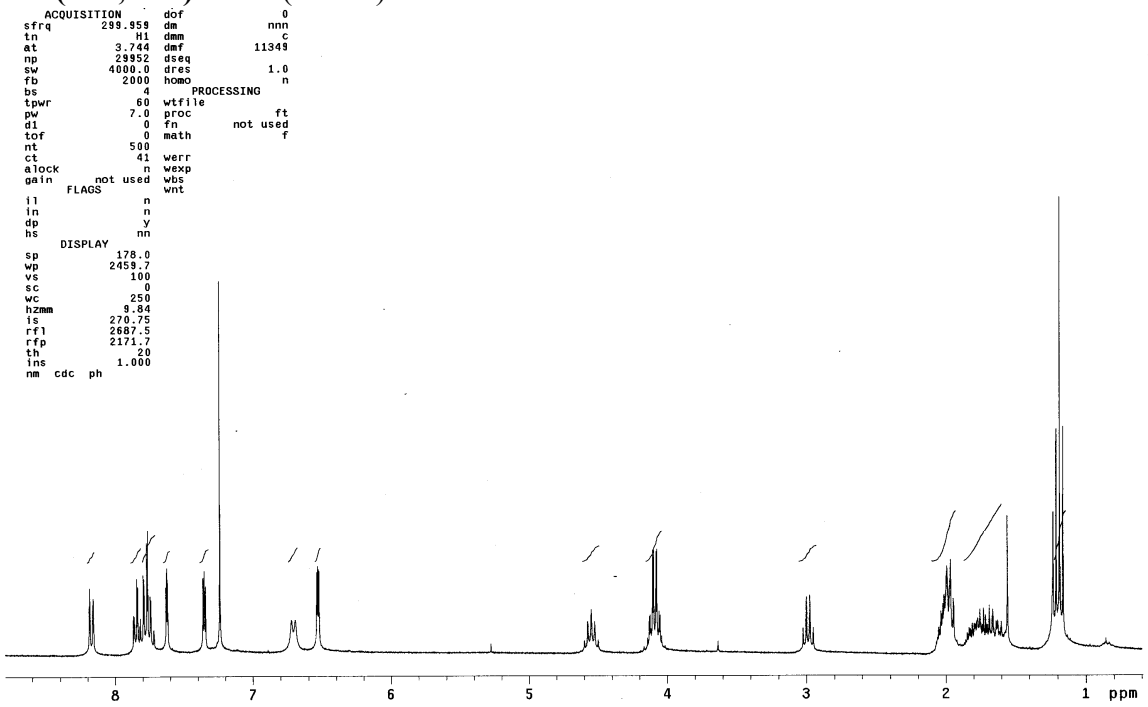
22k

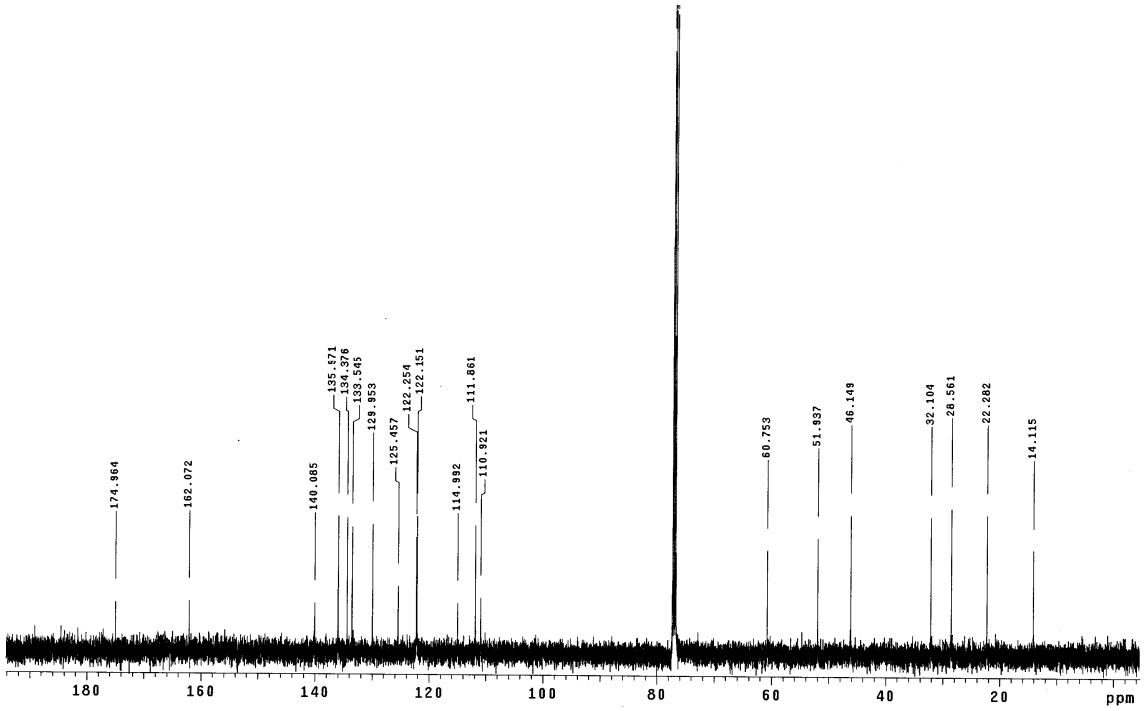
Compound **22k** was prepared from (25 mg, 0.1 mmol) of **20** and 2-cyanobenzenesulfonyl chloride. Flash chromatography (20 % to 50 % EtOAc/Hexanes) afforded 13.7 mg (33 %) **22k** as a white solid.

^{1}H NMR (300 MHz, $CDCl_3$) δ 8.20-8.15 (m, 1H), 7.88-7.61 (m, 3H), 7.64-7.61 (m, 1H), 7.33-7.38 (m, 1H), 6.71 (d, 1H, J = 8.4 Hz), 6.56-6.50 (m, 1H), 4.62-4.48 (m, 1H), 4.16-4.02 (m, 2H), 3.04-2.94 (m, 1H), 2.09-1.54 (m, 6H), 1.19 (t, 3H, J = 7.3 Hz)

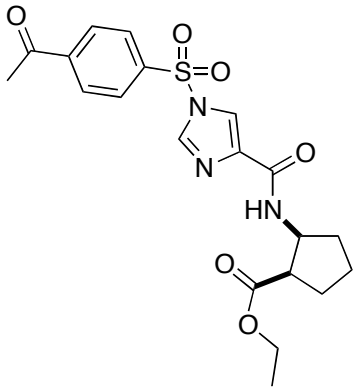
^{13}C NMR (125 MHz, $CDCl_3$) δ 175.0, 162.1, 140.1, 136.0, 134.4, 133.5, 130.0, 125.5, 122.3, 122.2, 115.0, 111.9, 110.9, 60.8, 51.9, 46.1, 32.1, 28.6, 22.3, 14.1

MS (ESI, m/z) 438.1 $(M+Na)^+$





General Procedure for Preparation of 23a-c. To a stirred mixture of 1*H*-imidazole-4-carboxylic acid (33.6 mg, 0.3 mmol), ethyl *cis*-2-amino-1-cyclopentane carboxylate hydrochloride (38.7 mg, 0.2 mmol) and HOBt (27.0 mg, 0.2 mmol) in 1.5 mL DMF were added EDC•HCl (38.2 mg, 0.2 mmol) and then *N*-methylmorpholine (33 μL, 0.3 mmol) under N_2 at 0 °C. The reaction mixture was stirred at room temperature and left stirring overnight, and then evaporated. The residue was dissolved in EtOAc and washed with water. The organic layer was washed successively with saturated NH_4Cl, saturated $NaHCO_3$, water and brine, and dried over anhydrous Na_2SO_4. Concentration of the organic layer afforded the crude amide **21** as a yellow oil. This product was used for the next reaction without further purification. To a solution of imidazole **21** (1 equiv) in anhydrous THF (0.10 M) was added Et_3N (2.0 equiv) in one portion. To this mixture was added dropwise over 10 min a solution of sulfonyl chloride in anhydrous THF (0.10 M), and the reaction mixture was left stirring overnight at room temperature. The precipitate of Et_3N•HCl was separated by filtration and the filtrate was evaporated *in vacuo*. Flash chromatography with EtOAc/hexanes mixtures provided pure compounds **23a-c**.



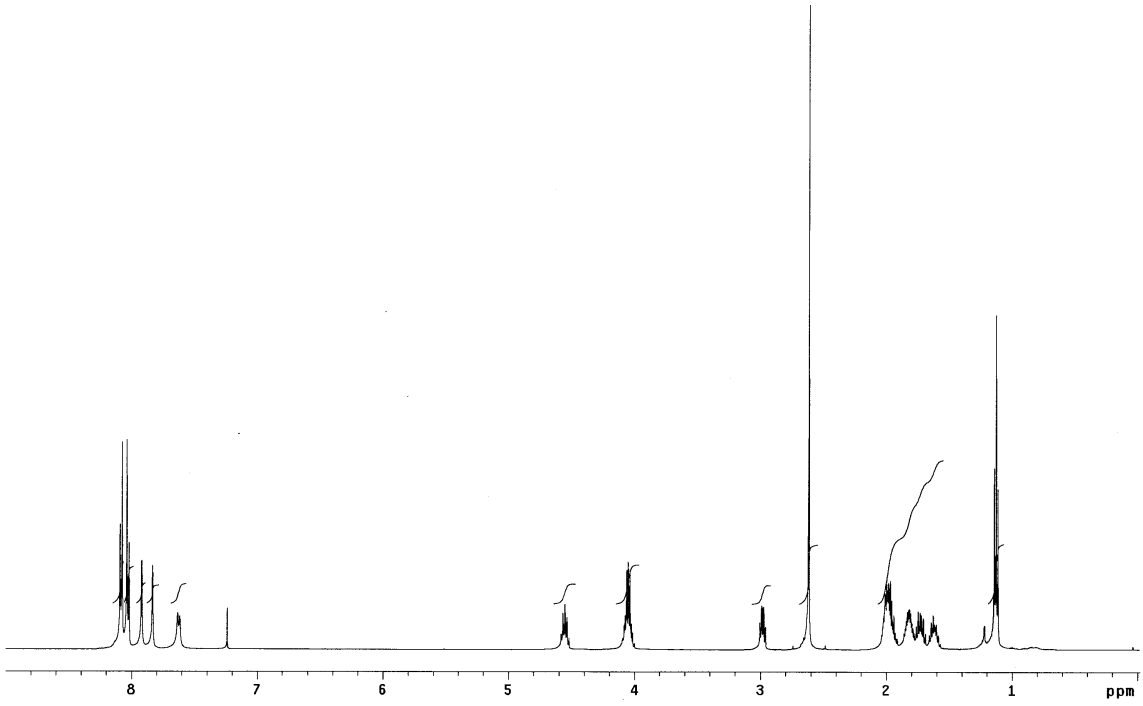
23a

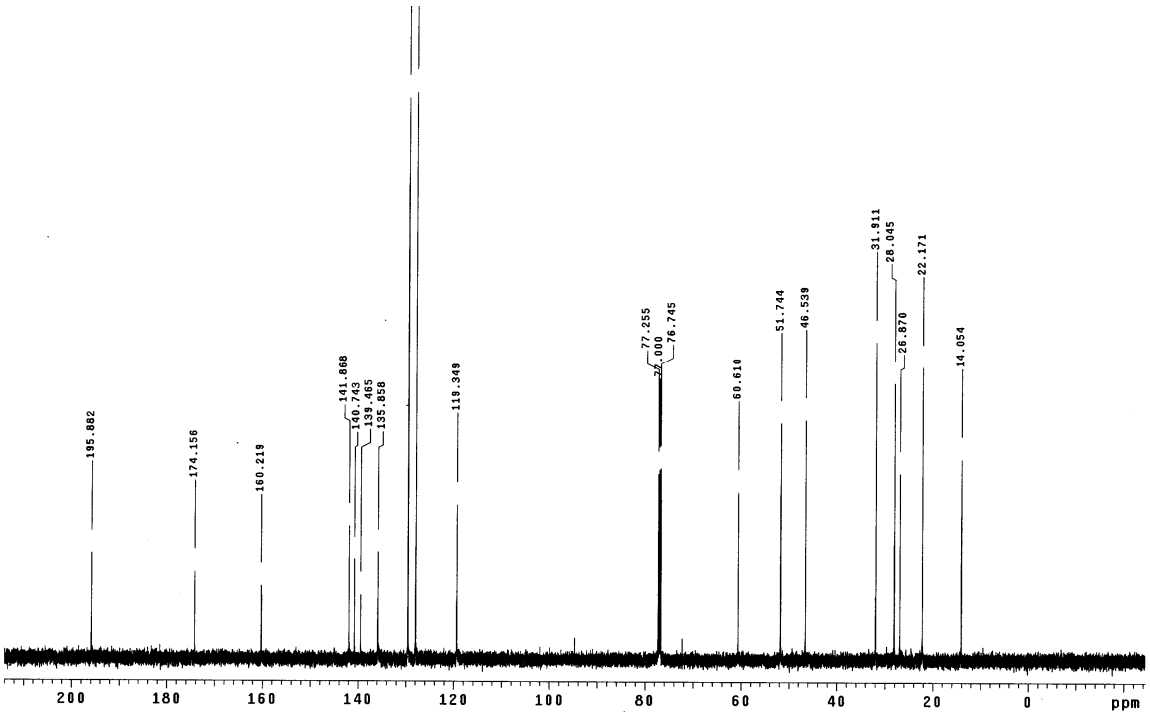
Compound **23a** was prepared from (19.4 mg, 0.1 mmol) of ethyl *cis*-2-amino-1-cyclopentane carboxylate hydrochloride and 1*H*-imidazole-4-carboxylic acid, and then 4-acetylbenzenesulfonyl chloride. Flash chromatography (20 % to 30 % EtOAc/Hexanes) afforded 29 mg (67 % in two steps) **23a** as a white solid.

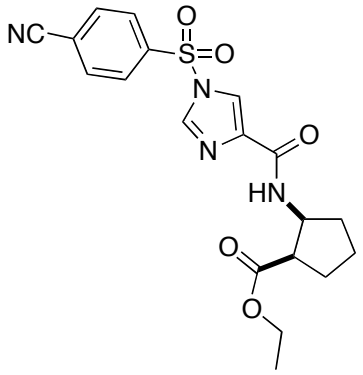
^{1}H NMR (500 MHz, $CDCl_3$) δ 8.08 (d, 2H, J = 9.5 Hz), 8.03 (d, 2H, J = 9.5 Hz), 7.92 (s, 1H), 7.83 (s, 1H), 7.62 (d, 1H, J = 9.0 Hz), 4.60-4.51 (m, 1H), 4.12-3.99 (m, 2H), 3.03-2.95 (m, 1H), 2.62 (s, 3H), 2.08-1.56 (m, 6H), 1.13 (t, 3H, J = 6.6 Hz)

^{13}C NMR (125 MHz, $CDCl_3$) δ 195.9, 174.2, 160.2, 141.9, 140.7, 139.5, 135.9, 129.6, 127.9, 119.3, 60.6, 51.7, 46.5, 31.9, 28.0, 26.9, 22.2, 14.1

MS (ESI, m/z) 434.1 $(M+H)^+$







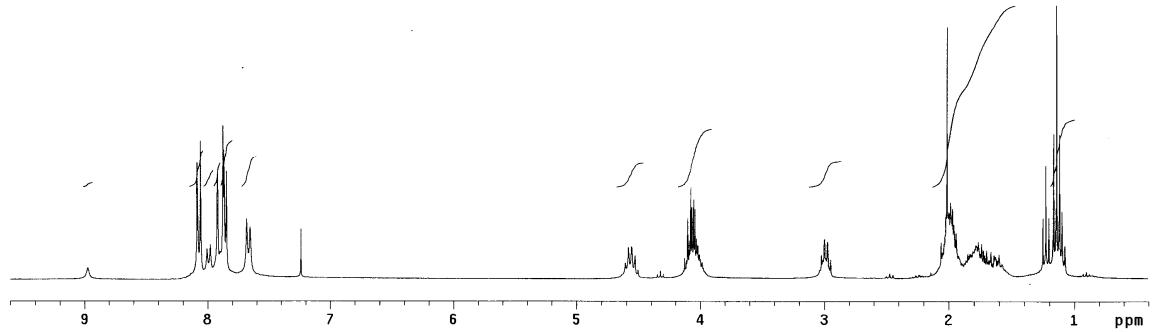
23b

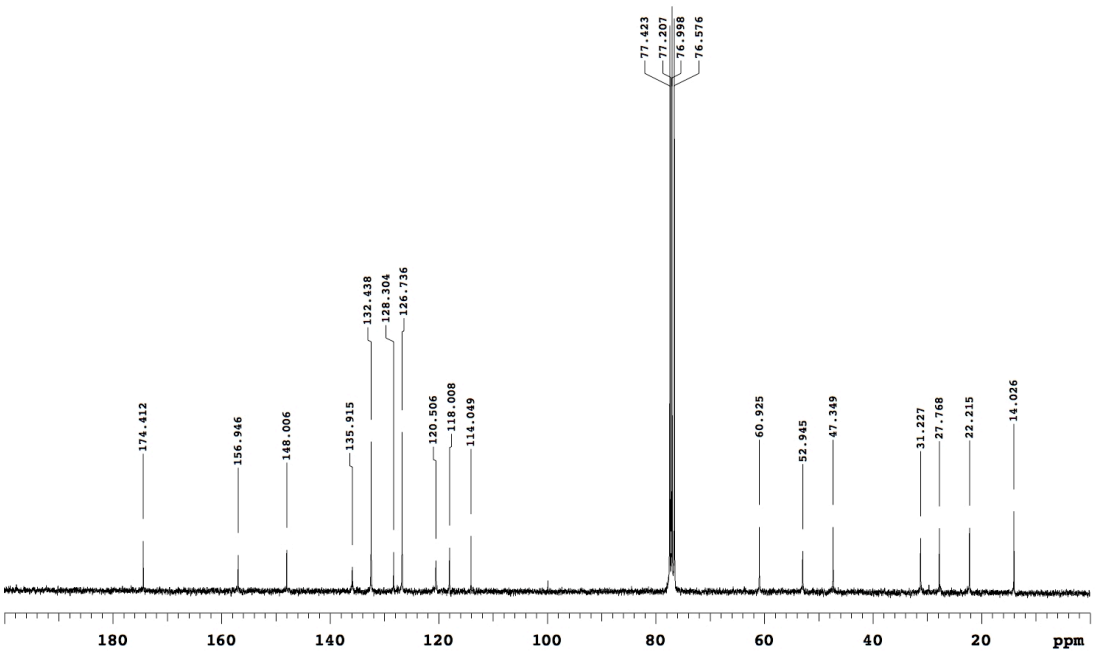
Compound **23b** was prepared from (19.4 mg, 0.1 mmol) of ethyl *cis*-2-amino-1-cyclopentane carboxylate hydrochloride and 1*H*-imidazole-4-carboxylic acid, and then 4-cyanobenzenesulfonyl chloride. Flash chromatography (20 % to 35 % EtOAc/Hexanes) afforded 27 mg (65 % in two steps) **23b** as a colorless oil.

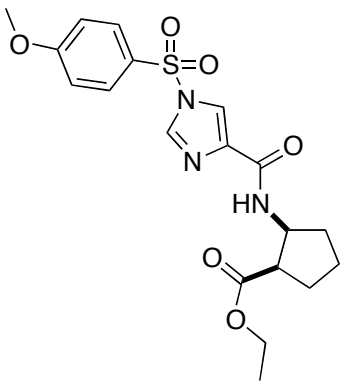
^{1}H NMR (500 MHz, $CDCl_3$) δ 8.07 (d, 2H, J = 9.0 Hz), 8.03-7.80 (m, 4H), 7.67 (d, 1H, J = 8.0 Hz), 4.66-4.48 (m, 1H), 4.18-3.82 (m, 2H), 3.07-2.93 (m, 1H), 2.10-1.50 (m, 6H), 1.13 (t, 3H, J = 6.9 Hz)

^{13}C NMR (75 MHz, $CDCl_3$) δ 174.4, 156.9, 148.0, 135.9, 132.4, 128.3, 126.7, 120.5, 118.0, 114.0, 60.9, 52.9, 47.4, 31.2, 27.8, 22.2, 14.0

MS (ESI, m/z) 439.1 $(M+Na)^+$







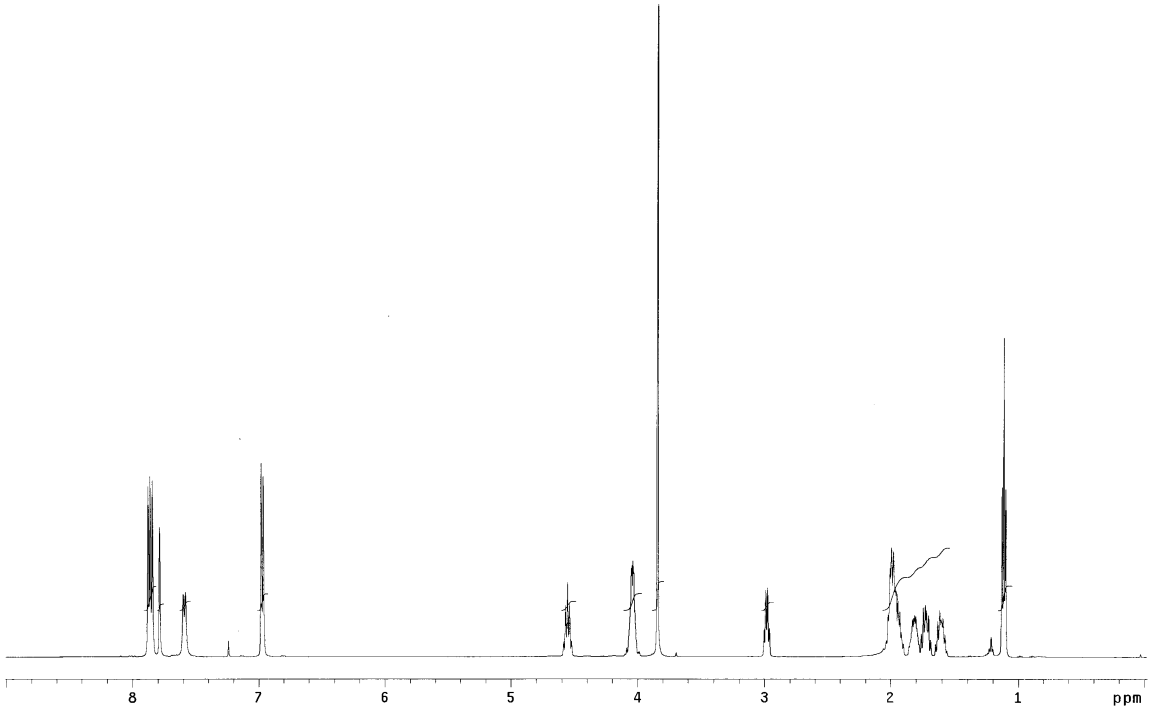
23c

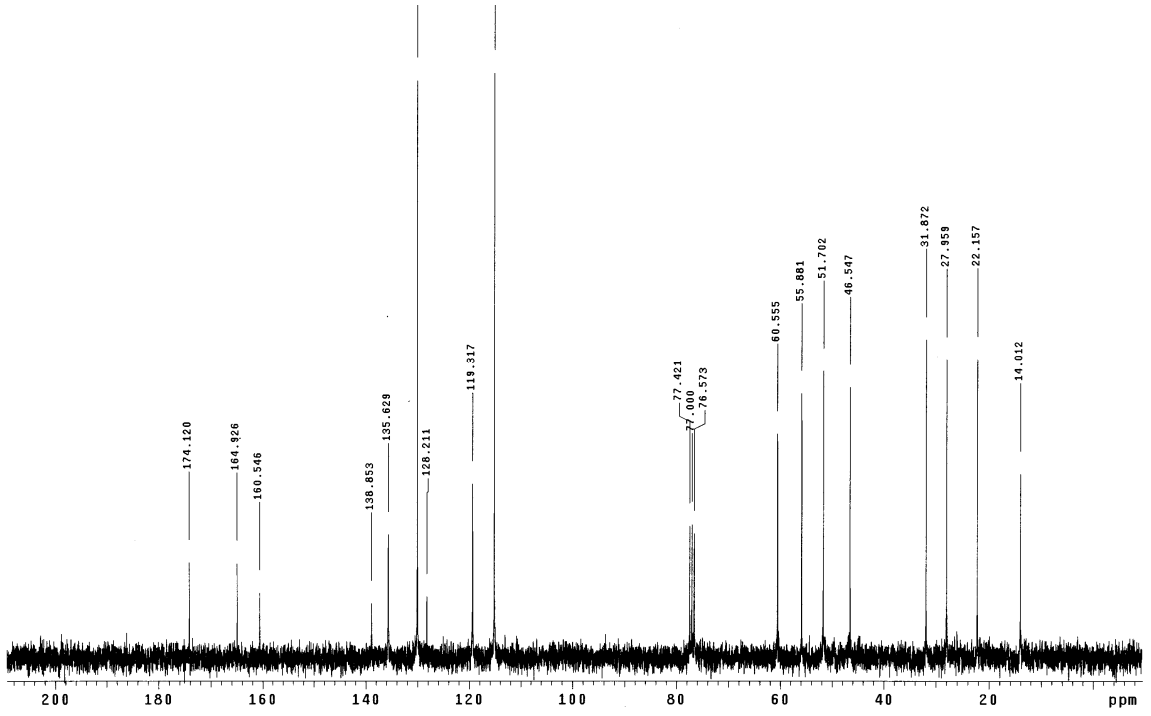
Compound **23c** was prepared from (19.4 mg, 0.1 mmol) of ethyl *cis*-2-amino-1-cyclopentane carboxylate hydrochloride and 1*H*-imidazole-4-carboxylic acid, and then 4-methoxybenzenesulfonyl chloride. Flash chromatography (20 % to 30 % EtOAc/Hexanes) afforded 33.7 mg (80 % in two steps) **23c** as a white solid.

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.90-7.83 (m, 3H), 7.78 (s, 1H), 7.59 (d, 1H, J = 9.0 Hz), 6.97 (d, 2H, J = 8.6 Hz), 4.60-4.51 (m, 1H), 4.10-3.98 (m, 2H), 3.84 (s, 3H), 3.02-2.95 (m, 1H), 2.06-1.55 (m, 6H), 1.12 (t, 3H, J = 6.6 Hz)

^{13}C NMR (75 MHz, $CDCl_3$) δ 174.1, 164.9, 160.5, 138.9, 135.6, 130.0, 128.2, 119.3, 115.1, 60.6, 55.9, 51.7, 46.5, 31.9, 28.0, 22.2, 14.0

MS (ESI, m/z) 422.1 $(M+H)^+$


8
7
6
5
4
3
2
1
ppm


174.120
164.926
160.546
138.853
135.629
128.211
119.317
77.421
77.000
76.573
60.555
55.881
51.702
46.547
31.872
27.959
22.157
14.012
200
180
160
140
120
100
80
60
40
20
ppm

VITA

Name: Jing Liu

Permanent Address:
Fucheng Road, 37-303
Haidian District
Beijing P.O. Box 840
Beijing 100840
P. R. China

Education:

2008
Ph.D. Chemistry
Texas A&M University
College Station, TX 77843
Advisor: Dr. Kevin Burgess

2001
B.S. Chemistry
Nankai University
Weijin Road #94
Tianjin 300071
P. R. China
Advisor: Dr. Ju Zuo